Draft Document for Review January 29, 2013 12:52 pmSG24-8068-00

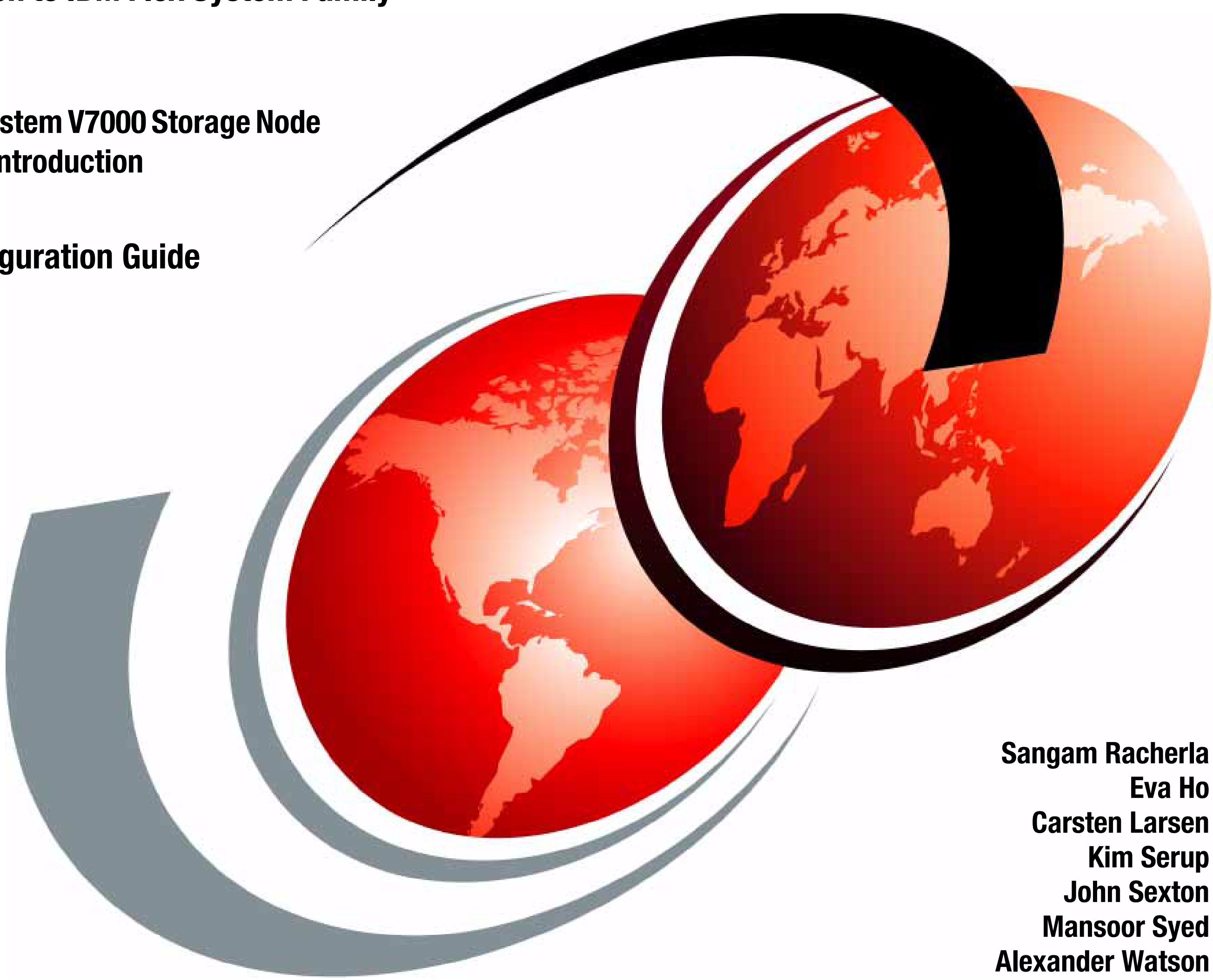
IBM Flex System V7000 Storage Node

Introduction and Implementation Guide

Introduction to IBM Flex System Family

IBM Flex System V7000 Storage Node Hardware Introduction

Host Configuration Guide



Sangam Racherla
Eva Ho
Carsten Larsen
Kim Serup
John Sexton
Mansoor Syed
Alexander Watson

Redbooks

ibm.com/redbooks

International Technical Support Organization

IBM Flex System V7000 Storage Node Introduction and Implementation Guide

June 2012

SG24-8068-00

Note: Before using this information and the product it supports, read the information in "Notices" on page ix.

First Edition (June 2012)

This edition applies to IBM Flex System V7000 Storage Node Version 6.4.

This document was created or updated on January 29, 2013.

Note: This book is based on a pre-GA version of a product and some of the features discussed in this book may change when the product becomes generally available. We recommend that you consult the product documentation or follow-on versions of this publication for more current information.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

Active Cloud Engine™
Active Memory™
AIX®
BladeCenter®
DS4000®
DS6000™
DS8000®
Easy Tier®
Electronic Service Agent™
EnergyScale™
FlashCopy®
HACMP™
IBM Flex System™
IBM Flex System Manager™
IBM®
Power Systems™
POWER7®
PowerVM®
POWER®
PureApplication™
PureFlex™
Real-time Compression™
Redbooks®
Redpapers™
Redbooks (logo) ®
ServerProven®
Storwize®
System i®
System p®
System Storage DS®
System Storage®
System x®
Tivoli®
X-Architecture®
XIV®
xSeries®

The following terms are trademarks of other companies:

Intel Xeon, Intel, Pentium, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

Preface

IBM® Flex System™ products are ideally suited for data center environments that require flexible, cost-effective, secure, and energy-efficient hardware.The IBM Flex System V7000 Storage Node is the latest addition to the IBM Flex Systems product family and is a modular storage system designed to fit into the IBM Flex System Enterprise chassis.

When purchased in the IBM PureFlex™ configurations, the IBM Flex System V7000 Storage Node is configured from the factory into the hardware solution purchased. If however, the desired configuration is not offered in the predefined offerings then a “Build to Order” configuration is designed to meet your needs.

The IBM Flex System V7000 Storage Node includes the capability to virtualize its own internal storage in the same manner as the IBM Storwize® V7000 does, and is designed to be a scalable internal storage system to support the compute nodes of the IBM Flex System environment.

This IBM Redbooks® publication introduces the features and functions of the IBM Flex System V7000 Storage Node through several examples. This book is aimed at pre- and post-sales technical support and marketing, storage administrators, and will help you understand the architecture of the IBM Flex System V7000 Storage Node, how to implement it, and take advantage of the industry leading functions and features.

The team who wrote this book

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, Raleigh Center.

Sangam Racherla is an IT Specialist and Project Leader working at the International Technical Support Organization (ITSO), San Jose Center. He holds a degree in Electronics and Communication Engineering and has eleven years of experience in the IT field. He has been with the ITSO for the past nine years and has extensive experience installing and supporting the ITSO lab equipment for various Redbooks publication projects. His areas of expertise include Microsoft Windows, Linux, IBM AIX®, IBM System x®, and IBM System p® servers, and various SAN and storage products.

Eva Ho is a worldwide Product Engineer support for IBM Flex System V7000.with IBM Systems Technology Group. She has 28 years of working experience within IBM, this includes: product development, L2/PFE support, and Product Engineer experience working with IBM products such as servers, networking products, IBM Network Attached Storage appliances, DS6000™, IBM System Storage® N series, Storwize V7000, Storwize V7000 Unified, and IBM Flex System V7000. She was also worked as technical team lead when she joined the STG worldwide N series PFE support team in Research Triangle Park, North Carolina. Eva has system storage certification with IBM. She was a participant in developing IBM Storage Networking Solutions V1 and V2 Certification test. Eva holds a master's degree in Computer Science.

Carsten Larsen is an IBM Certified Senior IT Specialist working for the Technical Services Support organization in IBM Denmark delivering consultancy services to IBM customers within the storage arena. Carsten joined IBM in 2007, leaving behind a job at HP where he worked with Storage Arrays and UNIX for 10 years, holding, among others, the HP

certification: Master Certified Systems Engineer (MASE). While working for IBM, Carsten has obtained Brocade BCFP and BCSD certifications as well as NetApp NCDA and NCIE certifications. Carsten is the author of a number of IBM Redbooks related to these product certifications.

Kim Serup is an IT Specialist at STG Lab Services, Copenhagen Denmark and has been with IBM since 1998. Kim is delivering consultancy services to IBM customers within the PureFlex, x Series and Storage arena. Before Kim joined STG Lab Services, he worked as a support specialist for IBM System x PFE support and provided EMEA Level 2 support for IBM DS3x00 products. At the beginning of Kim's career within IBM, Kim worked for the Technical Services Support organization, delivering hardware support and education for IBM customers, Business partners and IBM colleagues.

John Sexton is a Certified Consulting IT Specialist, based in Auckland, New Zealand, and has over 25 years of experience working in IT. He has worked at IBM for the last 16 years. His areas of expertise include IBM System p, AIX, HACMP™, virtualization, storage, IBM Tivoli® Storage Manager, SAN, SVC, and business continuity. He provides pre-sales support and technical services for clients throughout New Zealand, including consulting, solution design and implementation, troubleshooting, performance monitoring, system migration, and training. Prior to joining IBM in New Zealand, John worked in the United Kingdom supporting and maintaining systems in the financial and advertising industries.

Mansoor Syed is a Storage Support Specialist for the IBM Entry and Midrange Storage Systems, in Research Triangle Park, North Carolina. He provides hardware and software product support for IBM System Storage midrange and high performance computing systems as well as IBM virtualized disk system solutions. Mansoor joined IBM as a Client Advocate in 2004 and since has contributed in various technical support teams for products ranging from BladeCenter®, xSeries®, and System x servers to High Availability and High Performance Cluster Solutions. His responsibilities also involve enhancing RAS packaging of several IBM products, training of support engineers, and management of hardware infrastructure for Support teams. Presently, Mansoor provides technical leadership in the Product Field Engineering team for IBM System Storage DS5000 and DS3000 series.

Alexander Watson is an ATS Specialist for Storage Advanced Technical Skills (ATS) Americas in the United States. He is a Subject Matter Expert on the IBM Midrange system storage products and SAN switches. He has over fifteen years of experience in planning, managing, designing, implementing, problem analysis, and tuning of storage systems and SAN environments. He has worked at IBM for twelve years. His areas of expertise include IBM Midrange Storage solutions, SAN fabric networking, and Open System Storage IO performance. Al has been a co-author of more than thirteen Redbooks and Redpapers™.

Thanks to the following people for their contributions to this project:

Jon Tate
Tamikia Barrow

International Technical Support Organization, Raleigh Center

Matt Riddle
Mary Connell
Doris Konieczny
Roger Bullard
Walter Tita
Andy McNeil
Muhammad Zubair
Andy Sylivant

John Fasano
Karen Brown
Bill Wiegand
Andrew Martin
Dan Braden

IBM

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm
- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Introduction to IBM Flex Systems and IBM PureSystems offerings

This chapter provides an overview of the IBM PureSystems offerings and how the IBM Flex System V7000 Storage Node adds to a cloud ready solution within a single IBM Flex System Enterprise Chassis. Such a solution consists of Compute Nodes, Storage Systems and LAN- and SAN-infrastructure that connects it all together.

IBM Flex System products are ideally suited for data center environments that require flexible, cost-effective, secure, and energy-efficient hardware.

The innovative design features of the IBM Flex System products make it possible for you to configure totally integrated, customized, secure solutions that meet your data center needs today and provide flexible expansion capabilities for the future. The scalable hardware features and the unprecedented cooling capabilities of the IBM Flex System products help you optimize hardware utilization, minimize cost, and simplify the overall management of your data center.

The primary focus of this book is to describe features and functions of IBM Flex System V7000 Storage Node. However, in early versions of the IBM Flex System, the integrated storage is provided by IBM Storwize V7000. Hence this introduction covers both storage systems. In the chapters following mainly the IBM Flex System V7000 Storage Node will be discussed.

For more information of the IBM PureSystems please check:

http://www.ibm.com/ibm/puresystems/us/en/index.html

1.1 IBM PureSystems overview

During the last 100 years, information technology has moved from a specialized tool to a pervasive influence on nearly every aspect of life. From tabulating machines that simply counted with mechanical switches or vacuum tubes to the first programmable computers, IBM has been a part of this growth, while always helping customers to solve problems.

Information Technology (IT) is a constant part of business and of our lives. IBM expertise in delivering IT solutions has helped the planet become smarter. And as organizational leaders seek to extract more real value from their data, business processes and other key investments, IT is moving to the strategic center of business.

To meet those business demands, IBM is introducing a new category of systems—systems that combine the flexibility of general-purpose systems, the elasticity of cloud computing and the simplicity of an appliance that is tuned to the workload. Expert integrated systems are essentially the building blocks of capability. This new category of systems represents the collective knowledge of thousands of deployments, established best practices, innovative thinking, IT leadership and distilled expertise.

The offerings in IBM PureSystems are designed to deliver value in the following ways:

- Built-in expertise helps you to address complex business and operational tasks automatically.
- Integration by design helps you to tune systems for optimal performance and efficiency.
- Simplified experience, from design to purchase to maintenance, creates efficiencies quickly.

The IBM PureSystems offerings are optimized for performance and virtualized for efficiency. These systems offer a no-compromise design with system-level upgradeability. IBM PureSystems is built for cloud, containing “built-in” flexibility and simplicity.

At IBM, expert integrated systems come in two types:

- IBM PureFlex System. Infrastructure systems deeply integrate the IT elements and expertise of your system infrastructure.
- IBM PureApplication™ System. Platform systems include middleware and expertise for deploying and managing your application platforms

IBM PureSystems are built for cloud with integrated elasticity and virtualization capabilities to provision new services in minutes and improve business flexibility while reducing cost.

IBM Flex System is a build-to-order offering that is integrated by the client or a partner and does not deliver against all of the three attributes of expert integrated systems (built-in expertise, integration by design, simplified experience). IBM Flex System allows clients to build their own system to meet unique IT requirements with a set of no-compromise components including compute, storage, networking and systems management.

IBM PureFlex System and IBM PureApplication System are built on elements of the IBM Flex System, and it has been designed for clients that need pre-integrated hardware infrastructure comprised of compute, storage and networking nodes as well as a choice of operating systems and hypervisors.

The new IBM Flex System V7000 Storage Node is supported with IBM PureFlex System and other IBM Flex System configurations.

1.1.1 Product names

The primary product names for the IBM PureSystems components are as follows:

- IBM PureSystems
  - The overall name for IBMs new family of expert integrated systems.
- IBM Flex System
  - A build-to-order offering with clients choice of IBM Flex System components
  - IBM Flex System can help you go beyond blades
  - An innovative Enterprise Chassis designed for new levels of simplicity, flexibility, integration, reliability and upgradability
  - A broad range of x86 and IBM POWER® compute nodes
  - New IBM Flex System V7000 Storage Node built into the Enterprise Chassis
- IBM PureFlex System
  - A solution that combines compute nodes, storage, networking, virtualization and management into a single infrastructure system that is expert at sensing and anticipating resource needs to optimize your infrastructure.
- IBM PureApplication System
  - A platform system designed and tuned specifically for transactional web and database applications. Its workload-aware, flexible platform is designed to be easy to deploy, customize, safeguard and manage.
- IBM Flex System V7000 Storage Node
  - The product name for the IBM Flex System V7000 Storage Node family of controller and expansion enclosures. The IBM Flex System V7000 Storage Node is an add-on for the IBM Flex System Enterprise Chassis.
- IBM Flex System V7000 Control Enclosure
  - The controller enclosure of the IBM Flex System V7000 Storage Node. The IBM Flex System V7000 Control Enclosure is an add-on for the IBM Flex System Enterprise Chassis and mounts internally into this.
  - The IBM Flex System V7000 Control Enclosure provides 24 disk drive bays.
  - The IBM Flex System V7000 Control Enclosure supports block workloads only.
- IBM Flex System V7000 Expansion Enclosure
  - A SAS disk shelf with 24 disk drive bays that connects to the IBM Flex System V7000 Control Enclosure. The IBM Flex System V7000 Expansion Enclosure is an add-on for the IBM Flex System Enterprise Chassis and mounts internally into this.
- IBM Storwize V7000
  - The IBM Storwize V7000 is a disk system with built in IBM SAN Volume Controller (SVC) functionality that has the ability to virtualize a wide range of external storage systems from either IBM or other Storage vendors.
  - The IBM Storwize V7000 Control Enclosure provides choice of 24 x 2.5" Small Form Factor (SFF) disk drives or 12 x 3.5" Large Form Factor (LFF) disk drive form factors.
  - The IBM Storwize V7000 supports block workloads only.
- IBM Storwize V7000 Unified

  - IBM Storwize V7000 Unified is like the IBM Storwize V7000 a disk system that provides internal storage and external virtualization. However the IBM Storwize V7000 Unified also has file modules that provides NAS functionality like the CIFS and NFS protocols.
  - The Storwize V7000 Unified consolidates block and file workloads into a single system
- IBM Storwize V7000 Control Enclosure
  - The controller enclosure of the IBM Storwize V7000 storage system.
  - The IBM Storwize V7000 Control Enclosure provides 12 or 24 disk drive bays depending of the model
- IBM Storwize V7000 Expansion Enclosure
  - A SAS disk shelf with either 12 or 24 disk drive bays which can connect to either the IBM Storwize V7000 Control Enclosure or the IBM Flex System V7000 Control Enclosures.

Figure 1-2 on page 5 shows the different IBM PureSystems and their building blocks.

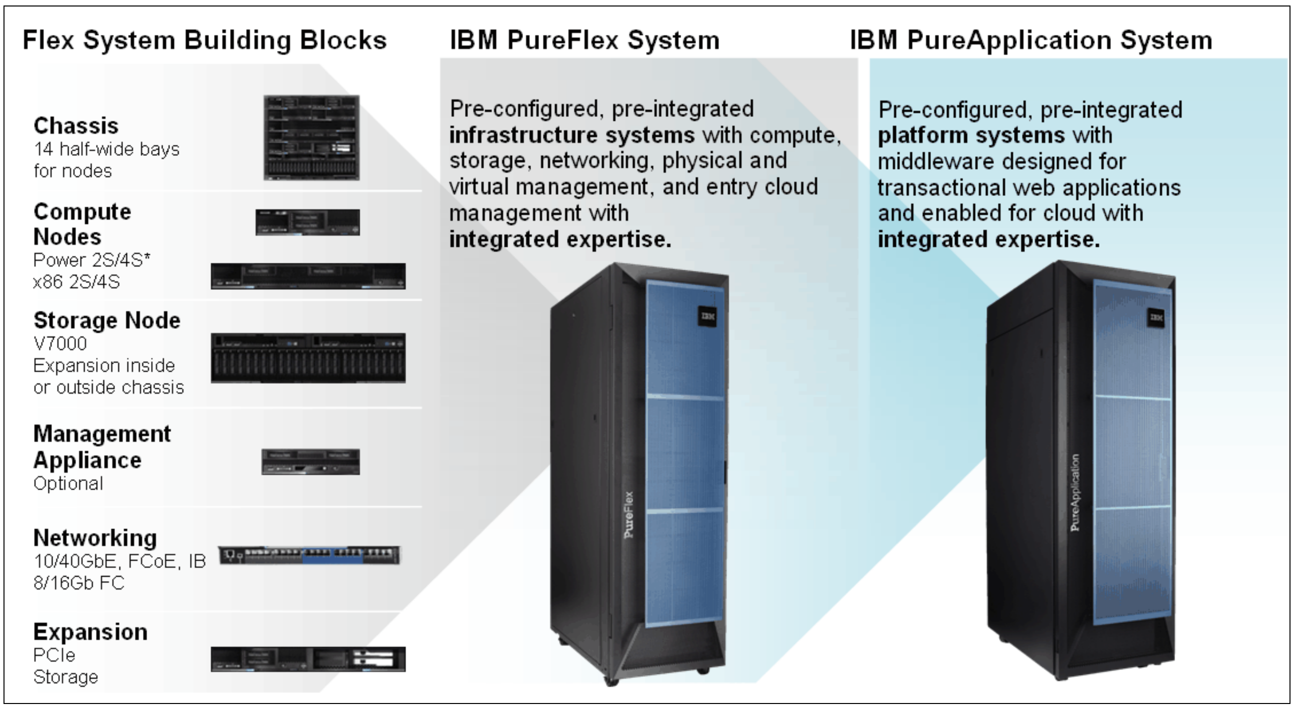


Figure 1-1 IBM PureSystems

1.1.2 IBM PureFlex System

To meet today's complex and ever-changing business demands, you need a solid foundation of server, storage, networking and software resources that is simple to deploy and can quickly and automatically adapt to changing conditions. You also need access to, and the ability to take advantage of, broad expertise and proven best practices in systems management, applications, hardware maintenance and more.

IBM PureFlex System is a comprehensive infrastructure system that provides an expert integrated computing system - combining servers, enterprise storage, networking, virtualization and management into a single structure. Its built-in expertise enables organizations to simply manage and flexibly deploy integrated patterns of virtual and hardware resources through unified management. These systems are ideally suited for

customers interested in a system that delivers the simplicity of an integrated solution but that also want control over tuning middleware and the run-time environment.

IBM PureFlex System recommends workload placement based on virtual machine compatibility and resource availability. Using built-in virtualization across servers, storage and networking, the infrastructure system enables automated scaling of resources and true workload mobility.

IBM PureFlex System undergoes significant testing and experimentation, so it can mitigate IT complexity without compromising the flexibility to tune systems to the tasks businesses demand. By providing both flexibility and simplicity, IBM PureFlex System can provide extraordinary levels of IT control, efficiency and operating agility that enable businesses to rapidly deploy IT services at a reduced cost. Moreover, the system is built on decades of expertise, enabling deep integration and central management of the comprehensive, open-choice infrastructure system and dramatically cutting down on the skills and training required for managing and deploying the system.

IBM PureFlex System combine advanced IBM hardware and software along with patterns of expertise and integrate them into three optimized configurations that are simple to acquire and deploy so you get fast time to value for your solution.

Figure 1-2 shows the IBM PureFlex System with its three different chassis implementations.

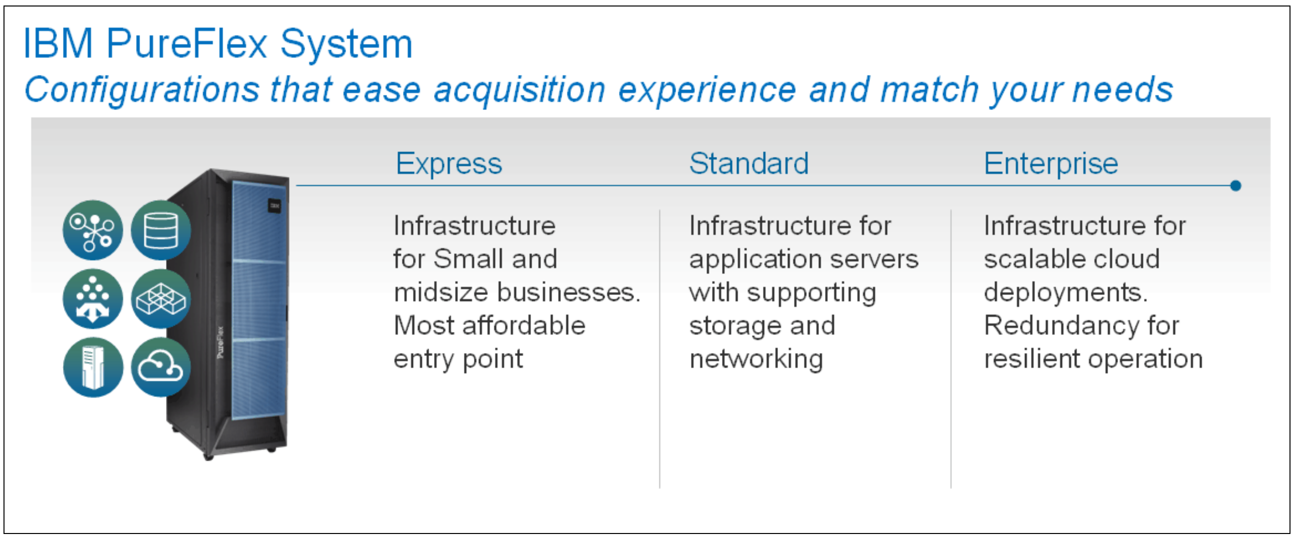


Figure 1-2 IBM PureFlex System

The three PureFlex System configurations are:

- IBM PureFlex System Express -- designed for small and medium businesses and is the most affordable entry point for PureFlex System
- IBM PureFlex System Standard -- optimized for application servers with supporting storage and networking and is designed to support your key ISV solutions
- IBM PureFlex System Enterprise -- optimized for transactional and database systems and has built-in redundancy for highly reliable and resilient operation to support your most critical workloads

> **Note:** IBM Flex System allows you to build your own system to meet the unique IT requirements.

The components of the PureFlex System are summarized in Table 1-1.

Table 1-1 IBM PureFlex System components

Component	IBM PureFlex System Express	IBM PureFlex System Standard	IBM PureFlex System Enterprise
IBM PureFlex System 42U Rack	1	1	1
IBM Flex System Enterprise Chassis	1	1	1
IBM Flex System Fabric EN4093 10Gb Scalable Switch	1	1	2 with both port-count upgrades
IBM Flex System FC3171 8Gb SAN Switch	1	2	2
IBM Flex System Manager Node	1	1	1
IBM Flex System Manager software license	IBM Flex System Manager with 1-year service and support	IBM Flex System Manager Advanced with 3-year service and support	IBM Flex System Manager Advanced with 3-year service and support
IBM Flex System Chassis Management Module	2	2	2
Chassis power supplies (std/max)	2 / 6	4 / 6	6 / 6
IBM Flex System Enterprise Chassis 80mm Fan Module Pair (std/max)	4 / 8	6 / 8	8 / 8
IBM Flex System V7000 Storage Node	Yes (redundant controller)	Yes (redundant controller)	Yes (redundant controller)
IBM Flex System V7000 Base Software	Base with 1-year software maintenance agreement	Base with 3-year software maintenance agreement	Base with 3-year software maintenance agreement

The fundamental building blocks of IBM PureFlex System solutions is the IBM Flex System Enterprise Chassis complete with compute nodes, networking and storage. Refer to the below sections for more information about the building blocks of the IBM PureFlex System.

- 1.2, “IBM PureFlex System Building Blocks” on page 8
- 1.3, “IBM Flex System Enterprise Chassis” on page 10
- 1.4, “Compute Nodes” on page 15
- 1.5, “I/O Modules” on page 23

1.1.3 IBM PureApplication System

IBM PureApplication System is a platform system that pre-integrates a full application platform set of middleware and expertise in with the IBM PureFlex System with a single management console. It is a workload-aware, flexible platform that is designed to be easy to deploy, customize, safeguard and manage in a traditional or private cloud environment, ultimately providing superior IT economics.

Note: IBM Flex System V7000 Storage Node is currently not offered in IBM PureApplication Systems. Currently the only available storage for IBM PureApplication Systems is IBM Storwize V7000.

With the IBM PureApplication System, you can provision your own patterns of software, middleware and virtual system resources. You can provision these patterns within a unique framework that is shaped by IT best practices and industry standards—standards that have been developed from many years of IBM experience with clients and from a deep understanding of smarter computing. These IT best practices and standards are infused throughout the system.

With IBM PureApplication System:

- IBM builds expertise into preintegrated deployment patterns, which can speed the development and delivery of new services.
- By automating key processes such as application deployment, PureApplication System built-in expertise capabilities can reduce the cost and time required to manage an infrastructure.
- Built-in application optimization expertise reduces the number of unplanned outages through best practices and automation of the manual processes identified as sources of those outages.
- Administrators can use built-in application elasticity to scale up or to scale down automatically. Systems can use data replication to increase availability.

Patterns of expertise can automatically balance, manage and optimize the elements necessary, from the underlying hardware resources up through the middleware and software. These patterns of expertise help deliver and manage business processes, services and applications by encapsulating best practices and expertise into a repeatable and deployable form. This best-practice knowledge and expertise has been gained from decades of optimizing the deployment and management of data centers, software infrastructures and applications around the world.

These patterns help you achieve the following types of value:

- Agility. As you seek to innovate to bring products and services to market faster, you need fast time-to-value. Expertise built into a solution can eliminate manual steps, automate delivery and support innovation.
- Efficiency. To reduce costs and conserve valuable resources, you must get the most out of your systems with energy efficiency, simple management and fast, automated response to problems. With built-in expertise, you can optimize your critical business applications and get the most out of your investments.
- Increased simplicity. You need a less complex environment. Patterns of expertise can help you easily consolidate diverse servers, storage and applications onto an easier-to-manage, integrated system.
- Control. With optimized patterns of expertise, you can accelerate cloud implementations to lower risk by improving security and reducing human error.

IBM PureApplication System is available in four configurations. These configuration options enable you to choose the size and compute power that meets your needs for application infrastructure. You can upgrade to the next size when your organization requires more capacity, and in most cases, you can do so without application downtime.

Table 1-2 on page 8 provides a high-level overview of the configurations.

Table 1-2 IBM PureApplication System configurations

	IBM PureApplication System W1500-96	IBM PureApplication System W1500-192	IBM PureApplication System W1500-384	IBM PureApplication System W1500-608
Cores	96	192	384	608
RAM	1.5 TB	3.1 TB	6.1 TB	9.7 TB
SSD Storage	6.4 TB			
HDD Storage	48.0 TB			
Application Services Entitlement	Included			

For more details about IBM PureApplication System, see

http://ibm.com/expert

1.2 IBM PureFlex System Building Blocks

IBM PureFlex System provides an integrated computing system - combining servers, enterprise storage, networking, virtualization, and management into a single structure. The built-in expertise lets organizations simply manage and flexibly deploy integrated patterns of virtual and hardware resources through unified management.

Each system consists of System x nodes, Power Systems™ compute nodes, or a combination of these two types, which is known as a hybrid configuration. The bundled, on-site services provide some initial compute node configuration and might differ for System x nodes and Power Systems compute nodes. A client-specified primary node (POWER or x86) is pre-configured with a hypervisor (PowerVM®, VMWare, KVM, HyperV) to allow virtual server configuration by IBM services personnel. Services also include skills transfer to the client personnel.

Note: Initial IBM PureFlex System configuration is carried out by IBM services and is included with the purchase. To ensure configuration success, the default shipped configuration must not be changed until these services are completed by IBM.

The IBM PureFlex System offerings comprise of the following components as illustrated in Figure 1-3 on page 9. With these components pre-configured, pre-integrated infrastructure systems with compute, storage, networking, physical and virtual management, and entry cloud management with integrated expertise are delivered by the IBM PureFlex System.

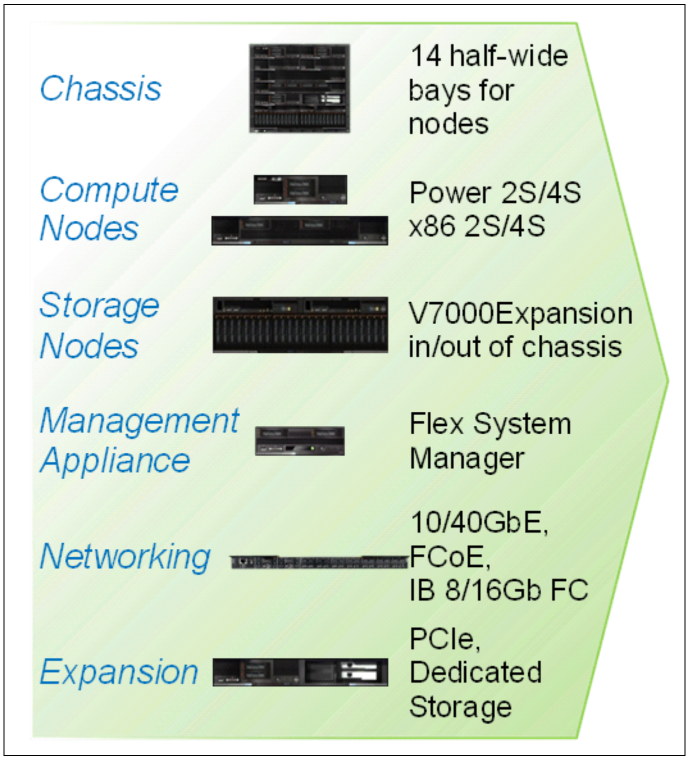


Figure 1-3 IBM PureFlex System Building Bocks

Storage components

The storage capabilities of IBM Flex System allows you to gain advanced functionality with storage nodes in your system while taking advantage of your existing storage infrastructure through advanced virtualization. For early versions of the IBM Flex System the only integrated storage was the IBM Storwize V7000 that was external to the IBM Flex System Enterprise Chassis. With the introduction of the IBM Flex System V7000 Storage Node storage is provided internally from the IBM Flex System Enterprise Chassis.

Simplified management

IBM Flex System simplifies storage administration with a single user interface for all your storage with a management console that is integrated with the comprehensive management system. These management and storage capabilities allow you to virtualize third-party storage with non-disruptive migration of the current storage infrastructure. You can also take advantage of intelligent tiering so you can balance performance and cost for your storage needs. The solution also supports local and remote replication and snapshots for flexible business continuity and disaster recovery capabilities.

Infrastructure

The IBM Flex System Enterprise Chassis is the foundation of the offering, supporting intelligent workload deployment and management for maximum business agility. The 10U high chassis has the capacity of up to 14 compute nodes, or a mix of compute nodes and Storage which mounts from the front. From the rear mounts power supplies, fans and different options of LAN and SAN switches. The IBM Flex System Enterprise Chassis delivers high-performance connectivity for your integrated compute nodes, storage, networking and management resources. The chassis is designed to support multiple generations of technology and offers independently scalable resource pools for higher utilization and lower cost per workload.

We now review the various components of the IBM Flex System in order to understand how the IBM Flex System V7000 Storage Node integrates with the PureFlex Systems solution. All of the components are used in the three pre-integrated offerings to support compute, storage, and networking requirements. You can select from these offerings, which are designed for key client initiatives and help simplify ordering and configuration.

While we only provide a summary of the IBM Flex System components in the following sections, its important to understand the various available options for IBM Flex System before we discuss the IBM Flex System V7000 Storage Node in detail in Chapter 2, “Introduction to IBM Flex System V7000 Storage Node” on page 39.

For detailed information about the components, refer to *IBM PureFlex System and IBM Flex System Products and Technology*, SG24-7984.

1.3 IBM Flex System Enterprise Chassis

The IBM Flex System Enterprise Chassis is a 10U next-generation server platform with integrated chassis management. It is a compact, high-density, high-performance, rack-mount, scalable server platform system. It supports up to 14 one-bay compute nodes that can share common resources, such as power, cooling, management, and I/O resources within a single Enterprise Chassis. In addition, it can also support up to seven 2-bay compute nodes or three 4-bay compute nodes (three IBM Flex System V7000 Storage Nodes or expansion enclosures) when the shelves are removed from the chassis. The1-bay, 2-bay, and 4-bay components can be ‘mixed and matched’ to meet specific hardware requirements.

Figure 1-4 on page 11 shows the chassis with the IBM Flex System V7000 Storage Node occupying 4 x compute bays which is partially inserted into the chassis for identification.

Figure 1-4 Front view of IBM Enterprise Flex System Chassis with an IBM Flex System V7000 Storage Node

The chassis has the following features on the front:

- The front information panel, located on the lower left of the chassis
- Bays 1 to 14 supporting Nodes, storage enclosures and FSM
- Lower airflow inlet apertures that provides air cooling for switches, IBM Flex System Chassis Management Module and power supplies
- Upper airflow inlet apertures that provide cooling for power supplies

For proper cooling, each bay in the front or rear in the chassis must contain either a device or a filler.

The Enterprise Chassis provides several LEDs on the front information panel that can be used to obtain the status of the chassis. The Identify, Check log and the Fault LED also appear on the rear of the chassis for ease of use.

The major components of Enterprise Chassis are:

- Fourteen 1-bay compute node bays (can also support seven 2-bay or three 4-bay compute nodes with shelves removed).
- Six 2500-watt power modules that provide N+N or N+1 redundant power.
- Ten fan modules (eight 80 mm fan modules and two 40 mm fan modules).

- Four physical I/O modules.
- An I/O architectural design capable of providing:
  - Up to 8 lanes of I/O to an I/O adapter card; each lane capable of up to 16 Gbps
  - A maximum of 16 lanes of I/O to a half wide-node with two adapters
  - A wide variety of networking solutions including Ethernet, Fibre Channel, FCoE, and InfiniBand
- Up to two IBM Flex System Manager (FSM) management appliances for redundancy. The FSM provides multiple-chassis management support for up to four chassis.
- Two IBM Flex System Chassis Management Module (CMMs). The CMM provides single-chassis management support.

The chassis can be configured with the following information about the chassis location:

- Rack Room
- Rack Location
- Position in Rack (the lowest Rack Unit occupied by the Chassis)
- Chassis Name
- Bay Id

Individual components will then be able to work out their bay in the chassis, the IBM Flex System V7000 Storage Node enclosure uses 4 bays (double wide and double tall) per enclosure and will report it's bay as the lowest left bay that it occupies.

Figure 1-5 shows the rear of the chassis where the IO modules and chassis management modules can be seen.

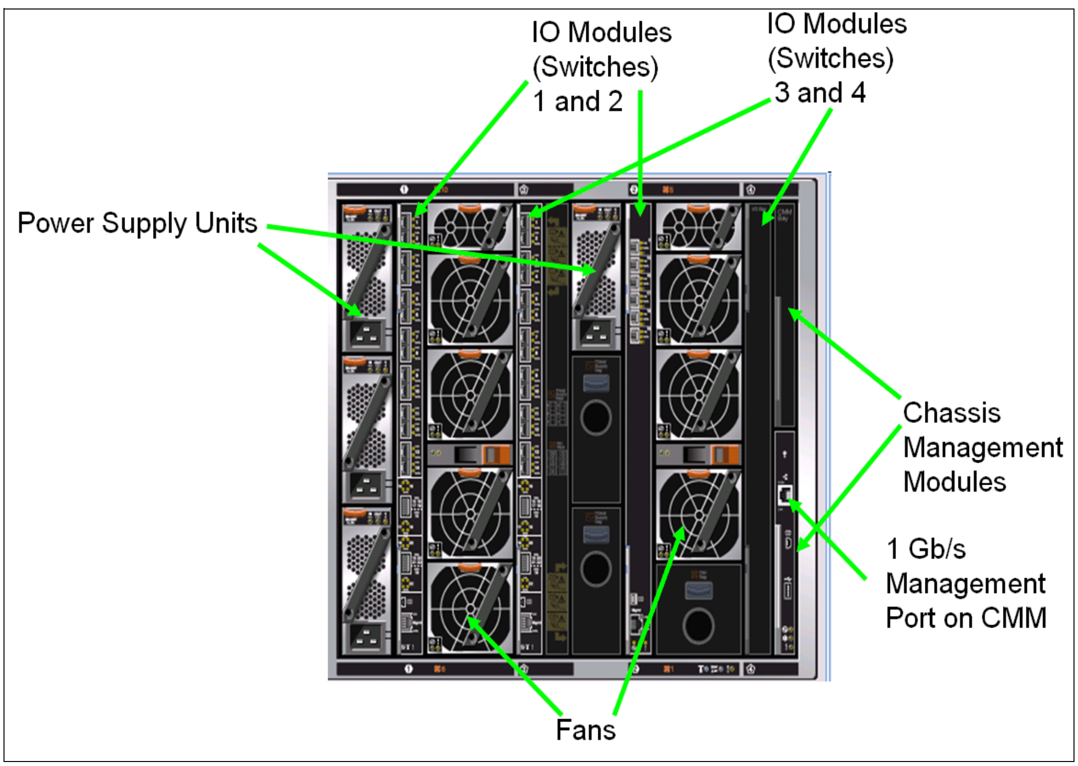


Figure 1-5 Rear view of the IBM Enterprise Flex System Chassis

The following components can be installed into the rear of the chassis

- Up to two IBM Flex System Chassis Management Modules (CMM).

- Up to six 2500W power supply modules.
- Up to six fan modules consisting of four 80 mm fan modules and two 40 mm fan modules. Additional fan modules may be installed to a total of ten modules.
- Up to four I/O modules.

1.3.1 Chassis Power supplies

A maximum of six power supplies may be installed within the Enterprise Chassis, the PSUs and empty PSU bays can be seen in Figure 1-5 on page 12. The power supplies are 80 PLUS Platinum certified and are 2100 Watts and 2500 Watts output rated at 200VAC, with oversubscription to 2895 Watts and 3538 Watts output at 200VAC respectively. The power supply operating range is 200-240VAC. The power supplies also contain two independently powered 40mm cooling fan modules, that pick power up from the midplane, not from the power supply.

Note: The 2100W power supplies are only available via Configure to Order (CTO). For more information about the 2100W power supply refer to *IBM PureFlex System and IBM Flex System Products and Technology*, SG24-7984.

The chassis allows configurations of power supplies to give N+N or N+1 redundancy. A fully configured chassis will operate on just three 2500W power supplies with no redundancy, but N+1 or N+N is likely to be preferred. Three (or six with N+N redundancy) power supplies allows for a balanced 3-phase configuration.

All power supply modules are combined into a single power domain within the chassis, which distributes power to each of the compute nodes, I/O modules and ancillary components through the Enterprise Chassis midplane. The midplane is a highly reliable design with no active components. Each power supply is designed to provide fault isolation and is hot swappable.

There is power monitoring of both the DC and AC signals from the power supplies, which allows the IBM Flex System Chassis Management Module to accurately monitor these signals. The integral power supply fans are not dependent upon the power supply being functional, they operate and are powered independently from the midplane.

Each power supply in the chassis has a 16A C20 3 pin socket and may be fed by a C19 power cable, from a suitable supply.

The chassis power system is designed for efficiency using datacenter power consisting of 3 phase, 60A Delta 200 VAC (North America) or 3 phase 32A wye 380-415 VAC (international). The Chassis may also be fed from single phase 200-240 VAC supplies if required.

Power Supply Redundancy

There are different vendors that may vary slightly in the terminology when describing power supply unit (PSU) redundancy. In general '*N*' is the minimum number of PSUs required to keep the server operational, in this case to keep the populated IBM Flex System Enterprise Chassis operational. The term '*N+1*' is the minimum number of PSUs plus one. This is not the best option for redundancy, it is the equivalent of a 'hot spare' and will protect against PSU failure only.

The minimum number of PSUs required to keep the server or chassis operational duplicated or doubled is referred to as '*N+N*'. The PSUs are fully redundant where there will be an even number for minimum '*N+N*' support.

An IBM Flex System Enterprise Chassis would typically be connected to at least two power distribution units (PDU) in a computer room with the PSUs connected evenly across the PDUs. With '*N+N*' redundancy the IBM Flex System Enterprise Chassis in this configuration is also protected against PDU failure within the computer room.

1.3.2 Fan modules and cooling

The Enterprise Chassis supports up to ten hot pluggable fan modules consisting of two 40mm fan modules and eight 80mm fan modules.

A chassis can operate with a minimum of six hot-swap fan modules installed, consisting of four 80mm fan modules and two 40mm fan modules. The fan modules plug into the chassis and connect to the fan distribution cards. The 80mm fan modules may be added as required to support chassis cooling requirements.

The two 40 mm fan modules in fan bays 5 and 10 (top two) distribute airflow to the I/O modules and chassis management modules. These modules ship pre installed in the chassis.

Each 40mm fan module contains two 40mm fans internally, side by side.

The 80 mm fan modules distribute airflow to the compute nodes through the chassis from front to rear. Each 80mm fan module contains two 80mm fan modules, back to back at each end of the module, which are counter rotating.

Both fan modules have an EMC (electromagnetic compatibility) mesh screen on the rear internal face of the module. The design of this also has an additional benefit for the airflow, by providing a laminar flow through the screen, which reduces turbulence of the exhaust air and improves the efficiency of the overall fan assembly. Laminar flow is a smooth flow of air, sometimes called streamline flow. The opposite of a laminar flow is a turbulent flow. The design of the whole fan assembly, the fan blade design, the distance between and size of the fan modules together with the EMC mesh screen ensures a highly efficient fan design that provides the best cooling for lowest energy input.

The minimum number of 80mm fan modules is four. The maximum number of 80mm fan modules that may be installed is eight. When the modules are ordered as an option, they are supplied as a pair.

Environmental Specifications

The chassis is designed to operate in temperatures up to 40°c (104°F), in ASHRAE class A3 operating environments.

The airflow requirements for the Enterprise Chassis are from 270 CFM (cubic feet per minute) to a maximum of 1020 CFM.

Environmental Specifications are:

- Humidity, non-condensing: -12°C dew point (10.4°F) and 8% - 85% relative humidity
- Maximum dew point: 24°C (75°F)
- Maximum elevation: 3050 m (10.006 ft)
- Maximum rate of temperature change: 5°C/hr (41°F/hr)

Heat Output (approximate):

- Maximum configuration: potentially 12.9kW

The 12.9kW figure is only a potential maximum, where the most power hungry configuration is chosen and all power envelopes are maximum. For a more realistic figure, the IBM Power

Configurator tool may be made to establish specific power requirements for a given configuration.

Power configurator can be found at:

http://www.ibm.com/systems/x/hardware/configtools.html

1.4 Compute Nodes

The IBM Flex System portfolio of compute nodes include those with Intel Xeon processors or with IBM POWER7® processors. Depending on the compute node design, it may come in one of two different form factors:

- Half-wide node -- occupies one chassis bay, half the width of the chassis (approximately 215 mm or 8.5”).
- Full-wide node -- occupies two chassis bays side-by-side, the full width of the chassis (approximately 435 mm (17”).

The applications installed on the compute nodes may be running on an operating system run natively on a dedicated physical server or may be virtualized in a virtual machine managed by a hypervisor layer. Here we provide the summary of the compute nodes, and for further detailed information on these, refer to *IBM Flex System p260 and p460 Planning and Implementation Guide*, SG24-7989 and *IBM PureFlex System and IBM Flex System Products and Technology*, SG24-7984.

1.4.1 IBM Flex System x440 Compute Node

The IBM Flex System x440 Compute Node (machine type 7917) is a high-density four socket server, optimized for high-end virtualization, mainstream database deployments, memory-intensive and high performance environments.

IBM Flex System x440 Compute Node is a double-wide compute node providing scalability to support up to four Intel Xeon E5-4600 processors and the node’s width allows fora significant I/O capability.

Figure 1-6 on page 16 shows the front of the compute node, showing the location of the controls, LEDs, and connectors. The light path diagnostic panel is located on the upper edge of the front panel bezel, in the same place as the x220 and x240.

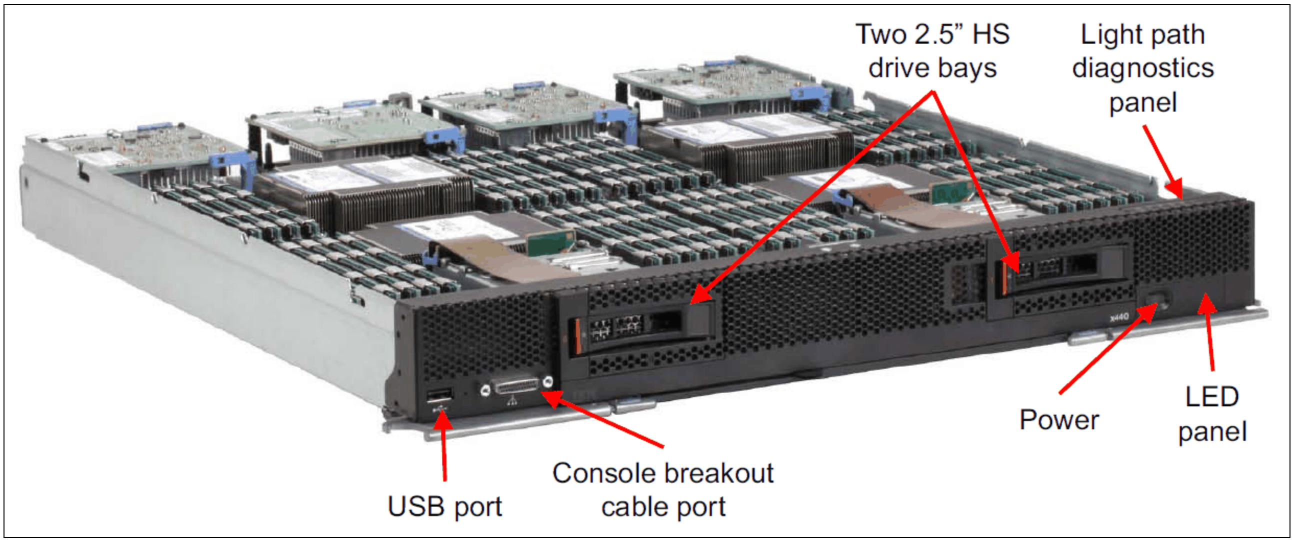


Figure 1-6 IBM Flex System x440 Compute Node

Refer to *IBM Flex System x440 Compute Node*, TIPS0886 for more information.

1.4.2 IBM Flex System x240 Compute Node

The IBM Flex System x240 Compute Node, available as machine type 8737 is a half-wide, two-socket server running the latest Intel Xeon processor E5-2600 family processors. It is ideal for infrastructure, virtualization and enterprise business applications and is compatible with the IBM Flex System Enterprise Chassis. The x240 supports up to two Intel Xeon E5-2600 series multi-core processors, 24 DIMM modules, two hot-swap drives, two PCI Express I/O adapter cards, and has an option for two internal USB connectors. Figure 1-7 shows the single bay x240 compute node.

Figure 1-7 The x240 type 8737

The IBM Flex System x240 Compute Node type 8737 features the Intel Xeon E5-2600 series processors with either two, four, six or eight cores per processor with up to 16 threads per socket. The processors have up to 20 MB of shared L3 cache, Hyper-Threading, Turbo Boost Technology 2.0 (depending on processor model), two QuickPath Interconnect (QPI) links that

run at up to 8 GT/s, one integrated memory controller, and four memory channels supporting up to three DIMMs each.

The x240 includes 8 GB of memory (2x 4 GB DIMMs) running at either 1600 MHz or 1333 MHz depending on model. Some models include an Embedded 10Gb Virtual Fabric Ethernet LAN-on-motherboard (LOM) controller as standard, this embedded controller precludes the use of an I/O adapter in I/O connector 1. Model numbers in the form x2x (for example 8737-L2x) include an Embedded 10Gb Virtual Fabric Ethernet LAN-on-motherboard (LOM) controller as standard. Model numbers in the form x1x (for example 8737-A1x) do not include this embedded controller.

The x240 with the Intel Xeon E5-2600 series processors can support up to 768 GB of memory in total when using 32 GB LRDIMMs and with both processors installed. The x240 uses Double Data Rate-3 (DDR-3) low-profile (LP) DIMMs. The x240 supports three types of DIMM memory:

- Registered DIMM (RDIMM) modules
- Unbuffered DIMM (UDIMM) modules
- Load-reduced (LRDIMM) modules

The mixing of these different memory DIMM types is not supported.

The x240 compute node features an onboard LSI 2004 SAS controller with two small form factor (SFF) hot-swap drive bays that are accessible from the front of the compute node. The onboard LSI SAS2004 controller provides RAID 0, RAID 1 or RAID 10 capability and supports up to two SFF hot-swap SAS or SATA HDDs or two SFF hot-swap solid state drives. Figure 1-8 shows how the LSI2004 SAS controller and hot-swap storage devices connect to the internal HDD interface.

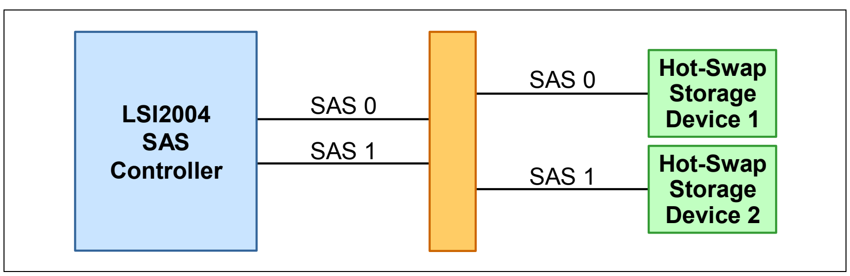


Figure 1-8 The LSI2004 SAS controller connections to HDD interface

Each x240 server has an Integrated Management Module version 2 (IMMv2) onboard and uses the Unified Extensible Firmware Interface (UEFI) to replace the older BIOS interface.

Embedded 10Gb Virtual Fabric Adapter

Some models of the x240 include an Embedded 10Gb Virtual Fabric Adapter (VFA, also known as LAN on Motherboard or LOM) built into the system board. Each of these models that includes the embedded 10Gb VFA also has the Compute Node Fabric Connector installed in I/O connector 1 (and physically screwed onto the system board) to provide connectivity to the Enterprise Chassis midplane.

I/O expansion

The x240 has two PCIe 3.0 x16 I/O expansion connectors for attaching I/O adapter cards. There is also another expansion connector designed for future expansion options. The I/O expansion connectors are a very high-density 216 pin PCIe connector. By installing I/O adapter cards it allows the x240 to connect with switch modules in the IBM Flex System Enterprise Chassis.

Figure 1-9 shows the rear of the x240 compute node and the locations of the I/O connectors.

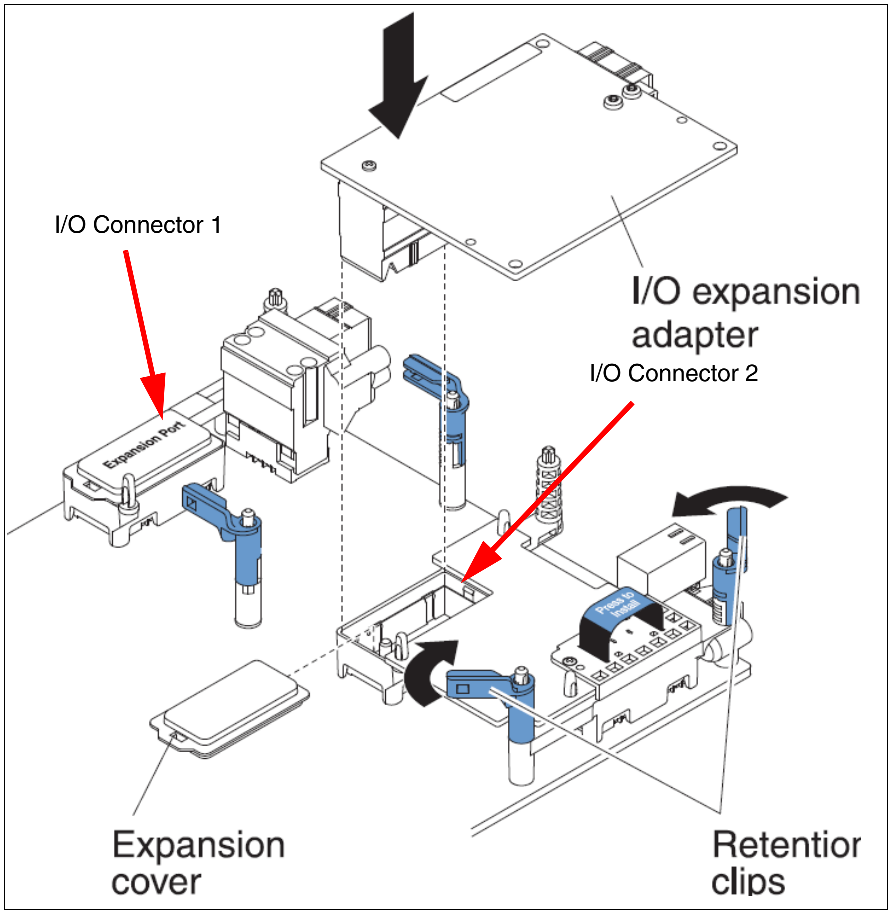


Figure 1-9 Rear of the x240 compute node showing the locations of the I/O connectors

1.4.3 IBM Flex System x220 Compute Node

The IBM Flex System x220 Compute Node, machine type 7906, is the next generation cost-optimized compute node designed for less demanding workloads and low-density virtualization. The x220 is efficient and equipped with flexible configuration options and advanced management to run a broad range of workloads. It is a high-availability, scalable compute node optimized to support the next-generation microprocessor technology. With a balance between cost and system features, the x220 is an ideal platform for general business workloads. This section describes the key features of the server.

Figure 1-10 on page 19 shows the front of the compute node indicating location of the controls, LEDs, and connectors.

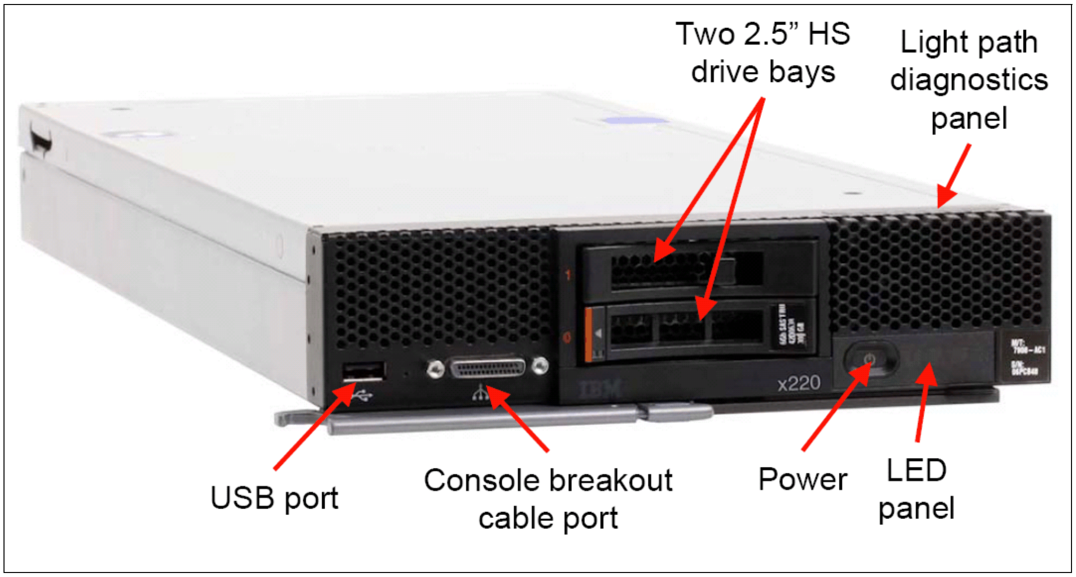


Figure 1-10 IBM Flex System x220 Compute Node

The IBM Flex System x220 Compute Node features the Intel Xeon E5-2400 series processors. The Xeon E5-2400 series processor has models with either four, six or eight cores per processor with up to 16 threads per socket. The processors have up to 20 MB of shared L3 cache, Hyper-Threading, Turbo Boost Technology 2.0 (depending on processor model), one QuickPath Interconnect (QPI) link that runs at up to 8 GT/s, one integrated memory controller, and three memory channels supporting up to two DIMMs each.

The x220 also supports an Intel Pentium 1403 or 1407 dual-core processor for entry-level server applications. Only one Pentium processor is supported in the x220. CPU socket 2 must be left unused and only six DIMM socks are available.

The x220 supports Low Profile (LP) DDR3 memory registered DIMMs (RDIMMs) and unbuffered DIMMs (UDIMMs). The server supports up to six DIMMs when one processor is installed and up to 12 DIMMs when two processors are installed. Each processor has three memory channels, and there are two DIMMs per channel.

The x220 server has two 2.5-inch hot-swap drive bays accessible from the front of the blade server as shown in Figure 1-10. The server optionally supports three internal disk controllers allowing a greater number of internal drives up to a maximum of eight with the ServeRAID M5115 controller and also supports 1.8-inch solid-state drives.

Each IBM Flex System x220 Compute Nodehas an Integrated Management Module version 2 (IMMv2) onboard and uses the Unified Extensible Firmware Interface (UEFI).

Embedded 1 Gb Ethernet controller

Some models of the x220 include an Embedded 1 Gb Ethernet controller (also known as LAN on Motherboard or LOM) built into the system board. Each x220 model that includes the controller also has the Compute Node Fabric Connector installed in I/O connector 1 (and physically screwed onto the system board) to provide connectivity to the Enterprise Chassis midplane.

The Fabric Connector enables port 1 on the controller to be routed to I/O module bay 1 and port 2 to be routed to I/O module bay 2. The Fabric Connector can be unscrewed and removed, if required, to allow the installation of an I/O adapter on I/O connector 1.

1.4.4 IBM Flex System p260 and p24L Compute Nodes

The IBM Flex System p260 Compute Node and IBM Flex System p24L Compute Node are based on IBM POWER architecture technologies. These compute nodes run in IBM Flex System Enterprise Chassis units to provide a high-density, high-performance compute node environment, using advanced processing technology. The IBM Flex System p24L Compute Node shares several similarities to the IBM Flex System p260 Compute Node in that it is a half-wide, Power Systems compute node with two POWER7 processor sockets,16 memory slots, two I/O adapter slots, and an option for up to two internal drives for local storage. The IBM Flex System p24L Compute Node is optimized for lower-cost Linux installations.

The IBM Flex System p260 Compute Node has the following features:

- Two processors with up to 16 POWER7 processing cores, up to 8 per processor
- Sixteen DDR3 memory DIMM slots supporting Active Memory™ Expansion
- Supports VLP (Very Low Profile) and LP (Low Profile) DIMMs
- Two P7IOC I/O hubs
- RAID-compatible SAS controller supporting up to 2 SSD or HDD drives
- Two I/O adapter slots
- Flexible Support Processor (FSP)
- System management alerts
- IBM Light Path Diagnostics
- USB 2.0 port
- IBM EnergyScale™ technology

The front panel of Power Systems compute nodes has the following common elements, as shown in Figure 1-11.

- USB 2.0 port
- Power-control button and light path, light-emitting diode (LED) (green)
- Location LED (blue)
- Information LED (amber)
- Fault LED (amber)

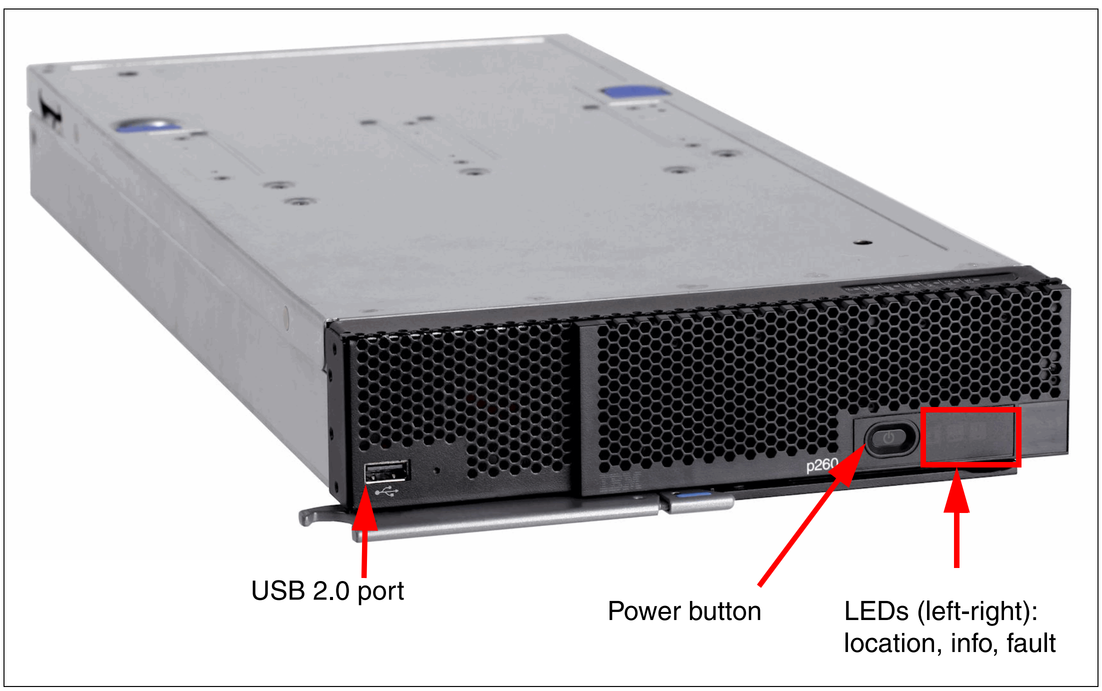


Figure 1-11 IBM Flex System p260 Compute Node with front panel details

There is no onboard video capability in the Power Systems compute nodes. The machines have been designed to be accessed using Serial Over LAN (SOL) or the IBM Flex System Manager (FSM).

The IBM POWER7 processor represents a leap forward in technology and associated computing capability. The multi-core architecture of the POWER7 processor has been matched with a wide range of related technologies to deliver leading throughput, efficiency, scalability, and reliability, availability, and serviceability (RAS).

Although the processor is an important component in servers, many elements and facilities have to be balanced across a server to deliver maximum throughput. As with previous generations of systems based on POWER processors, the design philosophy for POWER7 processor-based systems is one of system-wide balance in which the POWER7 processor plays an important role.

Each POWER7 processor has an integrated memory controller. Industry standard DDR3 Registered DIMM (RDIMM) technology is used to increase reliability, speed, and density of memory subsystems.

The p260 and p24L has an onboard SAS controller that can manage up to two, non-hot-pluggable internal drives. Both 2.5-inch hard disk drives (HDDs) and 1.8-inch solid-state drives (SSDs) are supported. The maximum number of drives that can be installed in the p260 or p24L is two. SSD and HDD drives cannot be mixed.

There are several advanced system management capabilities built into the p260 and p24L. A Flexible Service Processor handles most of the server-level system management. It has features, such as system alerts and Serial-Over-LAN capability.

A Flexible Service Processor (FSP) provides out-of-band system management capabilities, such as system control, run-time error detection, configuration, and diagnostics. Generally, you do not interact with the FSP directly but, rather, using tools, such as IBM Flex System Manager, IBM Flex System Chassis Management Module, and external IBM Systems Director Management Console. The FSP provides a Serial-over-LAN interface, which is available using the IBM Flex System Chassis Management Module and the **`console`** command.

The p260 and p24L do not have an on-board video chip and do not support keyboard, video, and mouse (KVM) connection. Server console access is obtained by a SOL connection only. SOL provides a means to manage servers remotely by using a command-line interface (CLI) over a Telnet or secure shell (SSH) connection. SOL is required to manage servers that do not have KVM support or that are attached to the IBM Flex System Manager. SOL provides console redirection for both System Management Services (SMS) and the server operating system. The SOL feature redirects server serial-connection data over a LAN without requiring special cabling by routing the data using the IBM Flex System Chassis Management Module network interface. The SOL connection enables Power Systems compute nodes to be managed from any remote location with network access to the IBM Flex System Chassis Management Module.

The IBM Flex System Chassis Management Module CLI provides access to the text-console command prompt on each server through a SOL connection, enabling the p260 and p24L to be managed from a remote location.

I/O adapter slots

There are two I/O adapter slots on the p260 and the p24L which are identical in shape (form factor). Also different is that the I/O adapters for the Power Systems compute nodes have their own connector that plugs into the IBM Flex System Enterprise Chassis midplane.

The I/O is controlled by two P7-IOC I/O controller hub chips. This provides additional flexibility when assigning resources within Virtual I/O Server (VIOS) to specific Virtual Machine/LPARs.

1.4.5 IBM Flex System p460 Compute Node

The IBM Flex System p460 Compute Node is also based on IBM POWER architecture technologies. This compute node is a full-wide, Power Systems compute node with four POWER7 processor sockets, 32 memory slots, four I/O adapter slots, and an option for up to two internal drives for local storage. It runs in IBM Flex System Enterprise Chassis units to provide a high-density, high-performance compute node environment, using advanced processing technology.

The IBM Flex System p460 Compute Node has the following features:

- Four processors with up to 32 POWER7 processing cores
- Thirty-two DDR3 memory DIMM slots that support IBM Active Memory Expansion
- Supports Very Low Profile (VLP) and Low Profile (LP) DIMMs
- Four P7IOC I/O hubs
- RAID-capable SAS controller that support up to two SSD or HDD drives
- Four I/O adapter slots
- Flexible Support Processor (FSP)
- System management alerts
- IBM Light Path Diagnostics
- USB 2.0 port
- IBM EnergyScale technology

The front panel of Power Systems compute nodes has the following common elements, as shown by the p460 in Figure 1-12:

- USB 2.0 port
- Power-control button and light path, light-emitting diode (LED) (green)
- Location LED (blue)
- Information LED (amber)
- Fault LED (amber)

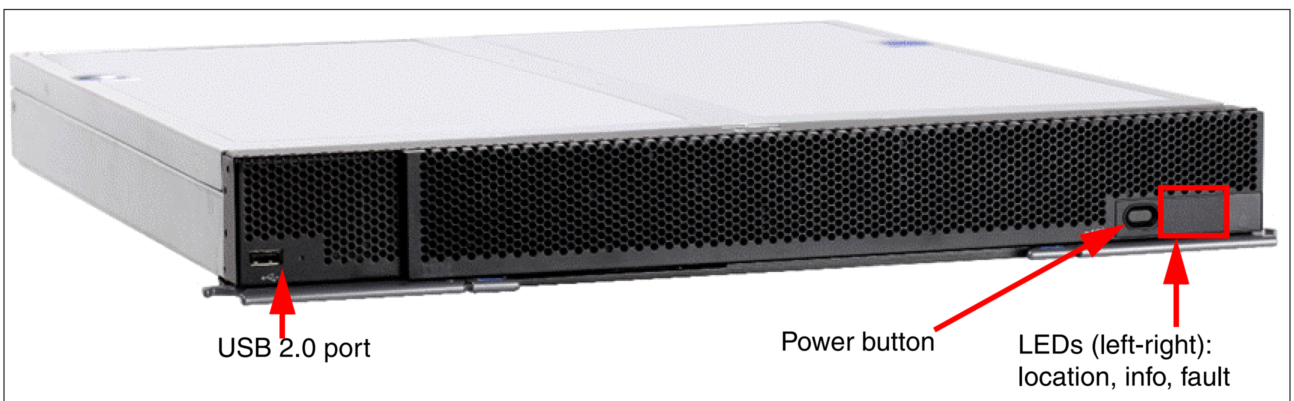


Figure 1-12 IBM Flex System p460 Compute Node showing front indicators

The USB port on the front of the Power Systems compute nodes is useful for a variety of tasks, including out-of-band diagnostics, hardware RAID setup, operating system access to data on removable media, and local OS installation. It might be helpful to obtain a USB optical (CD or DVD) drive for these purposes, in case the need arises as there is no optical drive in the IBM Flex System Enterprise Chassis

Although the processor is an important component in servers, many elements and facilities have to be balanced across a server to deliver maximum throughput. As with previous generations of systems based on POWER processors, the design philosophy for POWER7 processor-based systems is one of system-wide balance in which the POWER7 processor plays an important role.

Each POWER7 processor has two integrated memory controllers in the chip. Industry standard DDR3 Registered DIMM (RDIMM) technology is used to increase reliability, speed, and density of memory subsystems. The functional minimum memory configuration for the machine is 4GB (2x2GB) but that is not sufficient for reasonable production use of the machine. It is recommended for the IBM Flex System p460 Compute Node a minimum of 32GB of memory, with 32 x 16 GB DIMMs the maximum memory configurable is 512 GB.

The p460 has an onboard SAS controller that can manage up to two, non-hot-pluggable internal drives. Even though the p460 is a full-wide server, it has the same storage options as the p260 and the p24L.

The type of local drives used impacts the form factor of your memory DIMMs. If HDDs are chosen, then only very-low-profile (VLP) DIMMs can be used because of internal spacing. There is not enough room for the 2.5-inch drives to be used with low-profile (LP) DIMMs (currently the 2 GB and 16 GB sizes). Verify your memory choice to make sure it is compatible with the local storage configuration. The use of SSDs does not have the same limitation, and LP DIMMs can be used with SSDs.

The p460 System Management is the same as the p260 and p24L POWER compute nodes.

I/O adapter slots

The networking subsystem of the IBM Flex System Enterprise Chassis has been designed to provide increased bandwidth and flexibility. The new design also allows for more ports on the available expansion adapters, which will allow for greater flexibility and efficiency with your system's design.

There are four I/O adapter slots on the IBM Flex System p460 Compute Node. The I/O adapters for the p460 have their own connector that plugs into the IBM Flex System Enterprise Chassis midplane. There is no onboard network capability in the Power Systems compute nodes other than the Flexible Service Processor (FSP) NIC interface.

The I/O is controlled by four P7-IOC I/O controller hub chips. This provides additional flexibility when assigning resources within Virtual I/O Server (VIOS) to specific Virtual Machine/LPARs.

1.5 I/O Modules

The Enterprise Chassis can accommodate a total of four I/O modules which are installed in vertical orientation into the rear of the chassis, as shown in Figure 1-13 on page 24 where the four modules at the back of the chassis with the bays numbered. In addition to the two types of switches listed in Table 1-1 on page 6, there are alternative I/O modules that provide external connectivity, as well as connecting internally to each of the nodes within the chassis. They can be either Switch or Pass through modules with a potential to support other types in the future. These models can be ordered in 'build to order' IBM Flex Systems solutions.

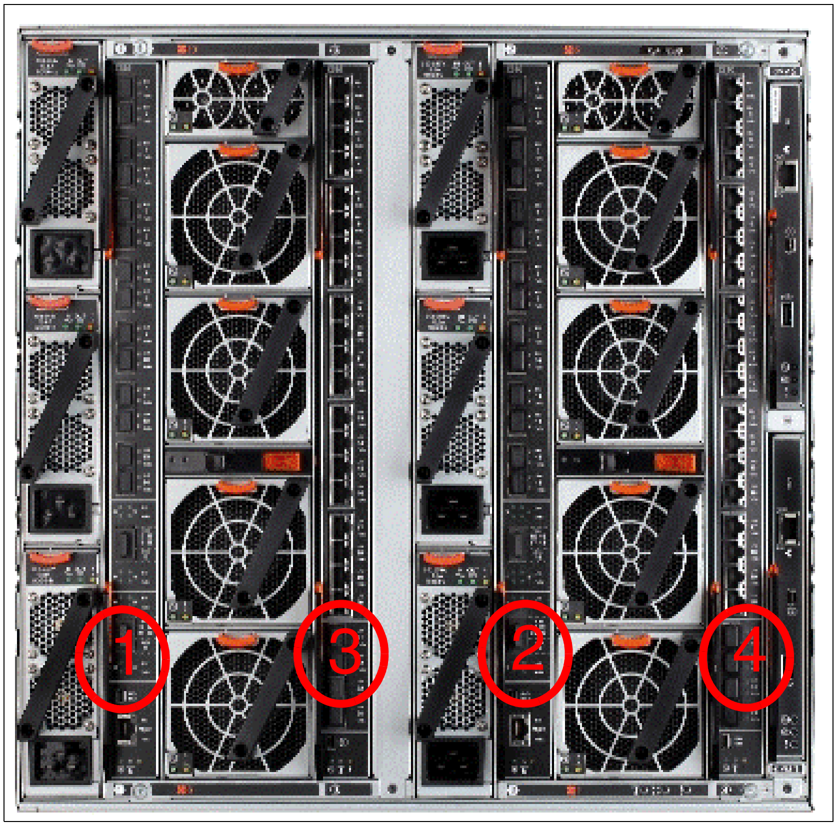


Figure 1-13 IBM Flex System Enterprise Chassis with I/O module bays numbered

If a node has a two port integrated LAN on motherboard (LOM) as standard, Module 1 and 2 are connected to this. If an I/O adapter is installed in the nodes I/O expansion bay 1, then Module 1 and 2 would be connected to this. Module 3 and 4 connect to the I/O adapter that is installed within I/O expansion bay 2 on the node.

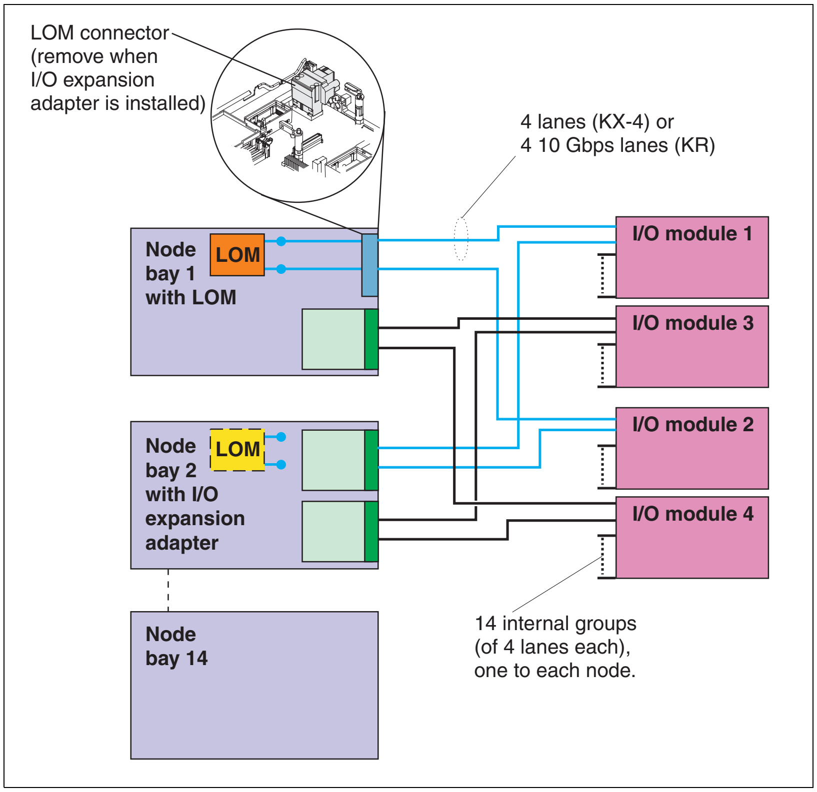


Figure 1-14 LOM, I/O adapter and switch module connection for node bays

The node in Bay 1 in Figure 1-14 shows that when shipped with a LOM, the LOM connector provides the link from the node motherboard to the midplane. Some nodes do not ship with LOM.

If required this LOM connector may be removed and an I/O expansion adapter installed in its place. This is shown on the node in Bay 2 in Figure 1-14.

1.5.1 IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch

The IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch provides support for L2 and L3 switching, Converged Enhanced Ethernet (PFC, ETS, DCBX), Fibre Channel over Ethernet (FCoE), NPV Gateway and Full Fabric Fibre Channel Forwarder (FCF).

The switch has the following major components:

- 42 10 Gb Ethernet internal ports and twenty-two external ports. External ports are arranged as two (small form-factor pluggable plus) SFP+ ports
- 12 SFP+ Omni Ports
- 2 Quad Small Form-Factor Pluggable Plus (QSFP+) ports.
- Each Omni Port is capable of running in 10 Gb Ethernet or 4/8 Gb FC mode with auto-negotiation capability.
- Support for Converged Enhanced Ethernet (CEE) and Fibre Channel over Ethernet (FCoE) over all Ethernet ports including Omni Ports (in Ethernet mode).
- Support for the Full Fabric FCF (Fibre Channel Forwarder) and NPV gateway.
- Support for full fabric FC services including Name Server, hardware-based Zoning.
- Support for IBM vNIC (virtual network interface card) Virtual Fabric Adapter with Single Root I/O Virtualization (SR-IOV) capability.

The 10 Gb Ethernet switch supports single compute node port capability (14 ports). Dual compute node port capability (28 ports) and triple compute node port capability (42 ports) are available with optional licenses.

The base model of this scalable switch provides:

- 14 internal 10 Gb Ethernet/FCoE ports
- 2 external 1Gb/10Gb Ethernet/FCoE ports
- 6 external flexible ports, usable for either 10 Gb Ethernet/FCoE or 4/8 Gb Fibre Channel.

With the optional licensing for pay-as-you-grow scalability, you easily and cost-effectively can enable additional internal 10 Gb Ethernet/FCoE ports, external 10 Gb/40 Gb Ethernet/FCoE ports and external flexible ports, usable for either 10 Gb Ethernet/FCoE or 4/8 Gb Fibre Channel.

For switch management access can be via:

  – A SSHv2/Telnet connection to the embedded command-line interface (CLI)
  – A terminal emulation program connection to the serial port interface
  – A Web browser-based interface (https/http) connection to the switch

1.5.2 IBM Flex System Fabric EN4093 and EN4093R 10Gb Scalable Switch

The IBM Flex System Fabric EN4093 and EN4093R 10Gb Scalable Switches are 10Gb 64-port upgradable midrange to high-end switch module, offering Layer 2/3 switching designed to install within the I/O module bays of the Enterprise Chassis. The switch has:

- Up to 42 internal 10Gb ports
- Up to 14 external 10Gb uplink ports (SFP+ connectors)
- Up to 2 external 40Gb uplink ports (QSFP+ connectors)

The switch is considered particularly suited for:

  – Building a 10Gb infrastructure
  – Implementing a virtualized environment
  – Investment protection for 40Gb uplinks
  – TCO reduction, improve performance, while maintaining high levels of availability and security
  – Oversubscription avoidance (Traffic from multiple internal ports attempting to pass through a lower quantity of external ports, leading to congestion & performance impact)

The rear of the switch has 14 SPF+ module ports and 2QSFP+ module ports. The QSFP+ ports can be used to provide either two 40Gb uplinks or eight 10Gb ports, using one of the supported QSFP+ to 4x 10Gb SFP+ cables. This cable splits a single 40Gb QSPFP port into 4 SFP+ 10Gb ports.

For management of the switch there is a mini USB port and also an Ethernet management port provided.

IBM Flex System Fabric EN4093R : IBM Flex System Fabric EN4093R's stacking capabilities simplify management for clients by stacking up to eight switches that share one IP address and one management interface. Support for Switch Partition (SPAR) allows clients to virtualize the switch with partitions that isolate communications for multitenancy environments.

For more information about the IBM Flex System Fabric EN4093 and EN4093R 10Gb Scalable Switches, refer to *IBM Flex System Fabric EN4093 and EN4093R 10Gb Scalable Switches*, TIPS0864.

1.5.3 IBM Flex System EN4091 10Gb Ethernet Pass-thru

The EN4091 10Gb Ethernet Pass-thru module offers a 1 for 1 connection between a single node bay and an I/O module uplink. It has no management interface and can support both 1 Gb and 10 Gb dual-port adapters installed in the compute nodes. If quad-port adapters are installed in the compute nodes, only the first two ports will have access to the pass-thru module's ports.

The necessary 1 GbE or 10 GbE module (SFP, SFP+ or DAC) must also be installed in the external ports of the pass-thru, to support the desired speed (1 Gb or 10 Gb) and medium (fiber optic or copper) for adapter ports on the compute nodes.

Four-port adapters: The EN4091 10Gb Ethernet Pass-thru has only 14 internal ports. As a result, only two ports on each compute node are enabled, one for each of two pass-thru modules installed in the chassis. If four-port adapters are installed in the compute nodes, ports 3 and 4 on those adapters are not enabled.

For more information about the IBM Flex System EN4091 10Gb Ethernet Pass-thru, refer to *IBM Flex System EN4091 10Gb Ethernet Pass-thru Module*, TIPS0865.

1.5.4 IBM Flex System EN2092 1Gb Ethernet Scalable Switch

The EN2092 1Gb Ethernet Switch provides support for L2/L3 switching and routing. The switch has:

- Up to 28 internal 1Gb ports
- Up to 20 external 1Gb ports (RJ45 connectors)
- Up to 4 external 10Gb uplink ports (SFP+ connectors)

The switch comes standard with 14 internal and 10 external Gigabit Ethernet ports enabled. Further ports can be enabled, including the four external 10Gb uplink ports.

For more information about the IBM Flex System EN2092 1Gb Ethernet Scalable Switch, refer to *IBM Flex System EN2092 1Gb Ethernet Scalable Switch*, TIPS0861.

1.5.5 IBM Flex System FC5022 16Gb SAN Scalable Switch

The IBM Flex System FC5022 16Gb SAN Scalable Switch is a high-density, 48-port 16 Gbps Fibre Channel switch that is used in the Enterprise Chassis. The switch provides 28 internal ports to compute nodes by way of the midplane, and 20 external SFP+ ports. These SAN switch modules deliver an embedded option for IBM Flex System users deploying storage area networks in their enterprise. They offer end-to-end 16 Gb and 8 Gb connectivity.

The N_Port Virtualization mode streamlines the infrastructure by reducing the number of domains to manage while enabling the ability to add or move servers without impact to the SAN. Monitoring is simplified via an integrated management appliance, or clients using end-to-end Brocade SAN can leverage the Brocade management tools.

Two versions are available, a 12-port switch module and a 24-port switch with the Enterprise Switch Bundle (ESB) software. The port count can be applied to internal or external ports using a a feature called Dynamic Ports on Demand (DPOD).

With DPOD, ports are licensed as they come online. With the FC5022 16Gb SAN Scalable Switch, the first 12 ports reporting (on a first-come, first-served basis) on boot-up are assigned licenses. These 12 ports may be any combination of external or internal Fibre Channel (FC) ports. After all licenses have been assigned, you can manually move those licenses from one port to another. As this is dynamic, no defined ports are reserved except ports 0 and 29. The FC5022 16Gb ESB Switch has the same behavior, the only difference is the number of ports.

For more information about the IBM Flex System FC5022 16Gb SAN Scalable Switch, refer to *IBM Flex System FC5022 16Gb SAN Scalable Switches*, TIPS0870.

1.5.6 IBM Flex System FC3171 8Gb SAN Switch

The IBM Flex System FC3171 8Gb SAN Switch is a full-fabric Fibre Channel switch module that can be converted to a pass-thru module when configured in transparent mode. This can be done using the switch GUI or CLI and then the Module can be converted back to a full function SAN switch at some future date. The switch requires a reset when turning transparent mode on or off.

The I/O module has 14 internal ports and 6 external ports. All ports are licensed on the switch as there are no port licensing requirements.

On this switch when in Full Fabric mode, access to all of the Fibre Channel Security features is provided. Security includes additional services available, being Secure Socket Layer (SSL) and Secure Shell (SSH). In addition, RADIUS servers may be used for device and user authentication. Once SSL/SSH is enabled, then the Security features are available to be configured. This allows the SAN administrator to configure which devices are allowed to login to the Full Fabric Switch module, by creating security sets with security groups. These are configured on a per switch basis. The security features are not available when in pass-thru mode.

The switch can be configured via either command line, or via QuickTools.

- Command Line: Access the switch via the console port through the IBM Flex System Chassis Management Module or through the Ethernet Port. This method requires a basic understanding of the CLI commands.
- QuickTools: Requires a current version of the JRE on your workstation before pointing a web browser to the switch's IP address. The IP Address of the switch must be configured. QuickTools does not require a license and code is included.

For more information about the IBM Flex System FC3171 8Gb SAN Switch, refer to *IBM Flex System FC3171 8Gb SAN Switch and Pass-thru*, TIPS0866.

1.5.7 IBM Flex System FC3171 8Gb SAN Pass-thru

The IBM Flex System FC3171 8Gb SAN Pass-thru I/O module is an 8 Gbps Fibre Channel pass-thru SAN module that has 14 internal ports and six external ports. It is shipped with all ports enabled.

Note: If there is a potential future requirement to enable full fabric capability, then this switch should not be purchased and instead the FC3171 8Gb SAN Switch should be considered.

The FC3171 8Gb SAN Pass-thru can be configured using either command line or QuickTools.

- Command Line: Access the module via the console port through the IBM Flex System Chassis Management Module or through the Ethernet Port. This method requires a basic understanding of the CLI commands.
- QuickTools: Requires a current version of the JRE on your workstation before pointing a web browser to the modules IP address. The IP Address of the module must be configured. QuickTools does not require a license and code is included.

For more information about the IBM Flex System FC3171 8Gb SAN Pass-thru, refer to *IBM Flex System FC3171 8Gb SAN Switch and Pass-thru*, TIPS0866.

1.5.8 IBM Flex System IB6131 InfiniBand Switch

IBM Flex System IB6131 InfiniBand Switch is a 32 port InfiniBand switch. It has 18 FDR/QDR (56/40 Gbps) external ports and 14 FDR/QDR (56/40 Gbps) internal ports for connections to nodes. This switch ships standard with QDR and can be upgraded to FDR.

Running the MLNX-OS, this switch has one external 1Gb management port and a mini USB Serial port for updating software and debug use, along with InfiniBand internal and external ports.

The switch has fourteen internal QDR links and eighteen CX4 uplink ports. All ports are enabled. The switch can be upgraded to FDR speed (56 Gbps) via the Feature On Demand (FOD) process.

> **Note:** Infiniband is not a supported protocol for IBM Flex System V7000 Storage Node or the Storwize V7000.

For more information about the IBM Flex System IB6131 InfiniBand Switch, refer to *IBM Flex System IB6131 InfiniBand Switch*, TIPS0871.

1.6 Introduction to IBM Flex System storage

Either the IBM Storwize V7000 or IBM Flex System V7000 Storage Node is an integrated part of the IBM PureFlex System depending on the model. Figure 1-15 on page 30 shows the IBM Flex System V7000 Storage Node where the left controller, called a cannister, is taken out. For the IBM Storwize V7000 products, the cannisters mounts from the rear where as in the IBM Flex System V7000 Storage Node, the controllers mounts from the front.

Figure 1-15 IBM Flex System V7000 Storage Node

The IBM Storwize V7000 product is described in detail in the *Implementing the IBM Storwize V7000 V6.3*, SG24-7938.

For more information about the IBM Flex System V7000 Storage Node see Chapter 2, “Introduction to IBM Flex System V7000 Storage Node” on page 39.

1.6.1 IBM Storwize V7000 and IBM Flex System V7000 Storage Node

IBM Storwize V7000 and IBM Flex System V7000 Storage Node are virtualized storage systems designed to complement virtualized server environments. They provide unmatched performance, availability, advanced functions and highly scalable capacity. IBM Storwize V7000 and IBM Flex System V7000 Storage Node are powerful disk systems that have been designed to be easy to use and enable rapid deployment without additional resources. IBM Storwize V7000 and IBM Flex System V7000 Storage Node supports block workloads whereas Storwize V7000 Unified (not covered in this book) consolidates block and file workloads into a single storage system for simplicity of management and reduced cost. IBM Storwize V7000 and IBM Flex System V7000 Storage Node supports IBM Real-time Compression™ for extraordinary efficiency, enabling storage of up to five times as much active data in the same physical disk space[1].

IBM Storwize V7000 and IBM Flex System V7000 Storage Node offer greater efficiency and flexibility through built-in solid state drive (SSD) optimization and thin provisioning technologies. IBM Storwize V7000 and IBM Flex System V7000 Storage Node advanced functions also enable non-disruptive migration of data from existing storage, simplifying implementation and minimizing disruption to users. Finally, these systems also enable you to virtualize and reuse existing disk systems, supporting a greater potential return on investment (ROI).

IBM Flex System V7000 Storage Node is designed to integrate into the IBM PureFlex System or IBM Flex System to enable extremely rapid storage deployment and breakthrough management simplicity. This new class of storage system combines no compromise design along with virtualization, efficiency, and performance capabilities of IBM Storwize V7000. It helps simplify and speed PureFlex System and IBM Flex System infrastructure deployment with superior server and storage management integration to automate and streamline

[1] Compression data based on IBM measurements. Compression rates vary by data type and content

provisioning and to help organizations achieve greater responsiveness to business needs while reducing costs.

For more information on IBM Real-time Compression in relation to IBM SAN Volume Controller and IBM Storwize V7000, refer to *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859.

Highlights

These are the highlights of IBM Storwize V7000 and IBM Flex System V7000 Storage Node:

- Delivers sophisticated enterprise-class storage function for businesses of all sizes
- Supports your growing business requirements while controlling costs
- Provides up to 200 percent performance improvement with automatic migration to high-performing Solid State Drives
- IBM Storwize V7000 and IBM Flex System V7000 Storage Node enables storing up to five times as much active data in the same physical disk space using IBM Real-time Compression[2]
- Enables near-continuous availability of applications through dynamic migration
- Supports faster and more efficient data copies for online backup, testing or data mining
- Offers flexible server and storage management with easy to use GUI for block and file storage management

IBM Storwize V7000 and IBM Flex System V7000 Storage Node are powerful block storage systems that combine hardware and software components to provide a single point of control to help support improved storage efficiency. By enabling virtualization, consolidation and tiering in business of all sizes, it is designed to improve application availability and resource utilization. The system offers easy-to-use, efficient and cost-effective management capabilities for both new and existing storage resources in your IT infrastructure.

Enhancing access with Easy Tier

Easy Tier® provides automatic migration of frequently accessed data to high performing Solid State Drives (SSDs), enhancing usage efficiencies. Operating at a fine granularity, the Easy Tier function automatically repositions pieces of the data to the appropriate class of drives based on I/O patterns and drive characteristics with no further administrative interaction.

Easy Tier makes it easy and economical to deploy SSDs in your environment. A hybrid pool of storage capacity is created containing two tiers: SSD and Hard Disk Drive (HDD). The busiest portions of volumes are identified and automatically relocated to high-performance SSDs. Remaining data can take advantage of higher capacity, price-optimized drives for the best customer value. Volumes in an SSD- or HDD-managed disk group are monitored and can be managed automatically or manually by moving hot extents to SSD and cold extents to HDD.

With an online database workload, Easy Tier improved throughput up to 200 percent and reduced transaction response time by up to 30 percent compared to a configuration using only HDD.[2]

Extraordinary storage efficiency

IBM Storwize V7000 and IBM Flex System V7000 Storage Node combines a variety of IBM technologies including thin provisioning, automated tiering, storage virtualization, Real-time Compression, clustering, replication, multi-protocol support and a next-generation graphical

[2] IBM lab measurements

user interface (GUI). Together, these technologies are designed to enable IBM Storwize V7000 and IBM Flex System V7000 Storage Node to deliver extraordinary levels of storage efficiency.

Newest of these technologies is IBM Real-time Compression, which is designed to improve efficiency by compressing data as much as 80 percent, enabling you to store up to five times as much data in the same physical disk space. Unlike other approaches to compression, IBM Real-time Compression is designed to be used with active primary data such as production databases and e-mail applications, which dramatically expands the range of candidate data that can benefit from compression. As its name implies, IBM Real-time Compression operates immediately as data is written to disk so no space is wasted storing uncompressed data awaiting post-processing.

The benefits of using IBM Real-time Compression together with other efficiency technologies are very significant and include reduced acquisition cost (because less hardware is required), reduced rack space, and lower power and cooling costs throughout the lifetime of the system. When combined with external storage virtualization, Real-time Compression can significantly enhance the usable capacity of your existing storage systems, extending their useful life even further.

IBM Real-time Compression is available for IBM Storwize V7000 and IBM Flex System V7000 Storage Node.

Avoiding disruptions with dynamic migration

IBM Storwize V7000 and IBM Flex System V7000 Storage Node use virtualization technology to help insulate host applications from physical storage changes. This ability can help enable applications to run without disruption while you make changes to your storage infrastructure. Your applications keep running so you can stay open for business.

Moving data is one of the most common causes of planned downtime. IBM Storwize V7000 and IBM Flex System V7000 Storage Node include a dynamic data migration function that is designed to move data from existing block storage into the new system or between arrays in a IBM Storwize V7000 and IBM Flex System V7000 Storage Node, while maintaining access to the data. The data migration function might be used, for example, when replacing older storage with newer storage, as part of load balancing work or when moving data in a tiered storage infrastructure.

Using the dynamic migration capabilities can provide efficiency and business value. Dynamic migration can speed time-to-value from weeks or months to days, minimize downtime for migration, eliminate the cost of add-on migration tools, and may help avoid penalties and additional maintenance charges for lease extensions. The result can be real cost savings to your business.

Foundation for cloud deployments

Improving efficiency and delivering a flexible, responsive IT infrastructure are essential requirements for any cloud deployment. Key technologies for delivering this infrastructure include virtualization, consolidation and automation.

With their virtualized storage design and tight affinity with technologies such as IBM PowerVM, and VMware, IBM Storwize V7000, IBM Flex System V7000 Storage Node, and IBM Storwize V7000 Unified are the ideal complement for virtualized servers that are at the heart of cloud deployments. IBM Storwize V7000 and IBM Flex System V7000 Storage Node help enable consolidation of multiple storage systems for greater efficiency. With IBM Storwize V7000 and IBM Flex System V7000 Storage Node, clustered systems drive the value of consolidation even further, and IBM Real-time Compression helps improve cost

effectiveness even further. Automated tiering technologies such as Easy Tier, Active Cloud Engine™ and Tivoli software help make the best use of the storage resources available.

Protecting data with replication services

IBM Storwize V7000 and IBM Flex System V7000 Storage Node supports block data while Storwize V7000 Unified supports both file and block data in the same system with replication functions optimized for the specific needs of each type of data.

Integrated management

IBM Storwize V7000 and IBM Flex System V7000 Storage Node provide a tiered approach to management designed to meet the diverse needs of different organizations. The systems' management interface is designed to give administrators intuitive control of these systems and provides a single integrated approach for managing both block and file storage requirements in the same system.

For organizations looking to manage both physical and virtual server infrastructures and the storage they consume (including provisioning and monitoring for higher availability, operational efficiency and infrastructure planning), IBM Storwize V7000, IBM Flex System V7000 Storage Node are integrated with IBM Systems Director Storage Control and IBM Flex System Manager™. A single administrator can manage and operate IBM servers, (IBM System x, IBM Power Systems, IBM BladeCenter and IBM PureFlex System) along with networking infrastructure and IBM storage from a single management screen.

High-performance SSD support

For applications that demand high disk speed and quick access to data, IBM provides support for SSDs in 200 and 400 GB 2.5-inch E-MLC (enterprise-grade multilevel cell) capacity. For ultra-high-performance requirements, IBM Storwize V7000 and IBM Flex System V7000 Storage Node may be configured with only SSDs for up to 96 TB of physical capacity in a single system (384 TB in a clustered system), enabling scale-out high performance SSD support.

External storage virtualization

External storage virtualization is the ability of IBM Storwize V7000 and IBM Flex System V7000 Storage Node to manage capacity in other disk systems. When IBM Storwize V7000 and IBM Flex System V7000 Storage Node virtualizes a disk system, its capacity becomes part of the IBM Storwize V7000 and IBM Flex System V7000 Storage Node systems and is managed in the same way as capacity on internal drives. Capacity in external disk systems inherits all the functional richness and ease-of-use of IBM Storwize V7000 and IBM Flex System V7000 Storage Node including advanced replication, thin provisioning, Real-time Compression and Easy Tier. Virtualizing external storage helps improve administrator productivity and boost storage utilization while also enhancing and extending the value of an existing storage asset.

> **External Storage Virtualization:** External Virtualization is only supported over FC or FCOE interface at this time. ISCSI is not supported for this.

For more information about External storage virtualization see Chapter 7, "Migration wizard" on page 265.

1.6.2 Benefits and value proposition

With IBM Storwize V7000 and IBM Flex System V7000 Storage Node you get the following Benefits:

- Simplified management and intuitive Graphical User Interface (GUI) - Aid rapid implementation and deployment.
- Virtualization of existing storage infrastructure improves administrator productivity.
- Improve space utilization up to 33 - 50 percent and up to 75 percent less capacity needed with FlashCopy® snapshots.
- With IBM Storwize V7000 and IBM Flex System V7000 Storage Node IBM Real-time Compression is designed to improve efficiency by storing up to five times as much active primary data in the same physical disk space. By significantly reducing storage requirements, you can keep up to five times more information online, use the improved efficiency to reduce storage costs, or achieve a combination of greater capacity and reduced cost.
- Increase storage performance up to 200 percent using Easy Tier technology.
- Dynamic migration helps decreasing migration times from weeks or months to days, eliminate the cost of add-on migration tools and provides continuous availability of applications by eliminating downtime.
- Thin provisioning allows you to purchase only the disk capacity needed.
- With IBM Storwize V7000 and IBM Flex System V7000 Storage Node clustered systems support the needs of growing business while enabling you to buy additional hardware only as needed.

Data Protection features

The following Data Protection features are supported with With IBM Storwize V7000 and IBM Flex System V7000 Storage Node:

- Volume Mirroring that allow a volume to remain online even when the storage pool backing it becomes inaccessible. The mirror copy is local to the system.
- Metro Mirror is a type of Remote Copy that creates a synchronous copy of data from a master volume to an auxiliary volume. The mirror copy is placed on a remote system and is an exact copy of the primary volume.
- Global Mirror provides an asynchronous copy, which means that the secondary volume is not an exact match of the primary volume at every point in time. The Global Mirror function provides the same function as Metro Mirror Remote Copy without requiring the hosts to wait for the full round-trip delay of the long-distance link. Like Metro Mirror the mirror copy is placed on a remote system.

> **Mirroring limitations:** Global (long distance of greater than 300 Km) mirroring is only supported over an FC SAN infrastructure. Local Mirroring (less than 300 Km) is supported over either an FC or FCoE SAN infrastructure. Mirroring over iSCSI fabric is not supported at this time. Consult System Storage Interoperability Center for supported configurations.

- Remote replication functions create exact copies of your data at remote locations to help you stay up and running in case of an emergency.
- FlashCopy and snapshot functions create instant copies of data to minimize data loss.

1.7 External Storage

In addition to the IBM Flex System V7000 Storage Node, the IBM Flex System Enterprise Chassis offers several possibilities for integration into storage infrastructure, such as Fibre Channel, iSCSI, and Converged Enhanced Ethernet

There are several options for attaching external storage systems to Enterprise Chassis, including:

- Storage area networks (SANs) based on Fibre Channel technologies
- SANs based on iSCSI
- Converged Networks based on 10 Gb Converged Enhanced Ethernet (CEE)

Fibre Channel-based SANs are the most common and advanced design of external storage infrastructure, they provide high levels of performance, availability, redundancy, and scalability. However, the cost of implementing FC SANs will be higher in comparison with CEE or iSCSI. The major components of almost every FC SAN include server's HBAs, FC switches, FC storage servers, FC tape devices, and optical cables for connecting these devices to each other.

iSCSI-based SANs provide all the benefits of centralized shared storage in terms of storage consolidation and adequate levels of performance, but use traditional IP-based Ethernet networks instead of expensive optical cabling. iSCSI SANs consist of server hardware iSCSI adapters or software iSCSI initiators, traditional network components such as switches, routers, and so forth, and storage servers with an iSCSI interface, such as IBM System Storage DS3500 or IBM N Series.

Converged Networks are capable of carrying both SAN and LAN types of traffic over the same physical infrastructure. Such consolidation allows you to decrease costs and increase efficiency in building, maintaining, operating, and managing of the networking infrastructure.

iSCSI, FC-based SANs, and Converged Networks can be used for diskless solutions to provide greater levels of utilization, availability, and cost effectiveness.

At the time of writing, the following IBM System Storage products are supported with the Enterprise Chassis:

- IBM Storwize V7000
- IBM XIV® Storage System series
- IBM System Storage DS8000® series
- IBM System Storage DS5000 series
- IBM System Storage DS3000 series
- IBM System Storage N series
- IBM System Storage TS3500 Tape Library
- IBM System Storage TS3310 Tape Library
- IBM System Storage TS3100 Tape Library

For the latest support matrices for storage products, refer to the storage vendors' interoperability guides. IBM storage products can be referenced in the IBM System Storage Interoperability Center (SSIC):

http://www.ibm.com/systems/support/storage/ssic/interoperability.wss

For the purpose of this book, we will limit the information to IBM Storwize V7000, and for more information about the other supported IBM System Storage products, refer to *IBM PureFlex System and IBM Flex System Products and Technology*, SG24-7984.

1.7.1 IBM Storwize V7000

IBM Storwize V7000 is an innovative storage offering that delivers essential storage efficiency technologies and exceptional ease of use and performance – all integrated into a compact, modular design. IBM Flex System V7000 Storage Node architecture is the same as that of the IBM Storwize V7000 and is managed from the IBM Flex System Chassis Management Module or IBM Flex System Manager node. IBM Storwize V7000 is considered external storage from a IBM PureFlex System perspective.

There are four levels of integration of Storwize V7000 with IBM Flex System as shown in Table 1-3

Table 1-3 Levels of integration

Level	Integration
Starting Level	IBM Flex System Single Point of Management
Higher Level	► Datacenter Management ► IBM Flex System Manager Storage Control
Detailed Level	► Data Management ► Storwize V7000 Storage User GUI
Upgrade Level	► Datacenter Productivity

For further information on IBM Storwize V7000, refer to *Implementing the IBM Storwize V7000 V6.3*, SG24-7938 and:

http://www.ibm.com/systems/storage/disk/storwize_v7000/overview.html

1.8 Useful websites

By the time this book is being written IBM Flex System V7000 Storage Node is pre-leased and support pages are not yet available. Function and features of the IBM Storwize V7000 and IBM Flex System V7000 Storage Node are almost identical. Some useful links for the IBM Storwize V7000 may be used as an alternative.

- IBM PureSystems entry pages:

  http://www.ibm.com/ibm/puresystems/us/en/index.html

- IBM PureSystems Redbooks pages:

  http://www.redbooks.ibm.com/portals/puresystems

- IBM Storwize V7000 and IBM Flex System Redbooks pages:

  http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=v7000

- IBM Flex System Information Center:

  http://publib.boulder.ibm.com/infocenter/flexsys/information/index.jsp

- IBM Flex System V7000 Storage Node Information Center:

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_product_page.html

- IBM Flex System V7000 Storage Node - Learning Modules

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.learning/site_learning_tutorials.html

- IBM Flex System V7000 Storage Node Support home:

  http://www.ibm.com/support/entry/portal/overview/hardware/puresystems/pureflex_system/storage_node/flex_system_v7000
- IBM Storwize V7000 Information Center:

  http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp
- The IBM Storwize V7000 Support page is at the following website:

  http://www.ibm.com/storage/support/storwize/v7000
- The IBM Storwize V7000 home page is at the following website:

  http://www-03.ibm.com/systems/storage/news/center/storwize_v7000/index.html
- The IBM Storwize V7000 Supported hardware list:

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003908
- The IBM Storwize V7000 Configuration Limit and Restrictions are at the following website:

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003902
- IBM Flex System V7000 Storage Node Installation Guide (GC27-4203)

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_bkmap_quickinstbk.pdf
- IBM Storwize V7000 Expansion Enclosure Installation Guide, Machine type 2076 (GC27-4234)

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/tbrd_bkmap_install_2076.pdf
- IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide (GC27-4205)

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_bkmap_trblshoot.pdf

2

Introduction to IBM Flex System V7000 Storage Node

In this chapter we introduce the IBM Flex System V7000 Storage Node and the enclosures and capabilities on which it is based. We will discuss in detail the controller and the expansion enclosures that make up the hardware of IBM Flex System V7000 Storage Node and point out the differences between them. We will present the concepts of virtualized storage, and how it works with the IBM Flex System V7000 Storage Node as well as briefly describe the many software features and capabilities that are available with this environment.

2.1 IBM Flex System V7000 Storage Node overview

When virtualizing external storage arrays, an IBM Flex System V7000 Storage Node can provide up to 32 PB of usable capacity. An IBM Flex System V7000 Storage Node supports a range of external disk systems similar to what the IBM Storwize V7000 system supports today. A control enclosure contains two control canisters; an expansion enclosure contains two expansion canisters. Both of these enclosures can contain up to 24 disk drives of the 2.5 inch form factor.

The IBM Flex System V7000 Storage Node is a modular storage system designed to fit into the IBM Flex System Enterprise chassis. When purchased in the IBM PureFlex configurations the IBM Flex System V7000 Storage Node is configured from the factory into the hardware solution purchased. If however, the desired configuration is not offered in the predefined offerings then a “Build to Order” configuration is designed to meet your needs.

The IBM Flex System V7000 Storage Node includes the capability to virtualize its own internal storage in the same manner as the IBM Storwize V7000 does. The IBM Flex System V7000 Storage Node is built upon the software base of the IBM Storwize V7000 which uses technology from IBM System Storage SAN Volume Controller (SVC) for virtualization and the advanced functions of the IBM System Storage DS8000 family for it RAID configurations of the internal disks, and the highly flexible graphical user interface (GUI) of the IBM XIV Storage Subsystem for management.

The IBM Flex System V7000 Storage Node provides a number of configuration options that are aimed at simplifying the implementation process. It also includes automated instruction steps, called *Directed Maintenance Procedures (DMP)*, to assist in resolving any events that might occur. The IBM Flex System V7000 Storage Node is a clusterable, scalable, storage system, and an external virtualization device.

The IBM Flex System V7000 Storage Node is designed to be a scalable internal storage system to support the compute nodes of the IBM Flex System environment. It will contain a control enclosure which contains a pair of clustered node canisters and accommodates up to twenty-four 2.5-inch disk drives within the enclosure. Each control enclosure can additionally attach a maximum of two IBM Flex System V7000 Expansion Enclosures which can reside in the IBM Flex System Enterprise chassis with it.

Optionally, up to nine IBM Storwize V7000 Expansion Enclosures can be installed externally; however, a total of no more than nine expansion enclosures using either IBM Flex System V7000 internal Expansion Enclosures (max 2), IBM Storwize V7000 external Expansion Enclosures (max 9), or any combination are supported. With the configuration of a single IBM Flex System V7000 Control Enclosure using twenty four 1TB 2.5" disks attached to nine IBM Storwize V7000 Expansion enclosures each using twelve 3TB 3.5" the system can manage a raw capacity of 348TB.

The IBM Flex System V7000 Storage Node can contain up to four control enclosures in a cluster configuration, each supporting the full configuration above, resulting in a maximum total raw capacity of 1392TB.

The control enclosure can support the addition of up to nine IBM Storwize V7000 expansion enclosures connected externally. The additional expansions allow for a total of 240 disks drives; or for a maximum capacity of 348TB supported per control enclosure.

Note: For maximum capacity numbers, the 7026-212 expansion enclosure and 3TB SAS Nearline drives must be used.

Figure 2-1 shows a representation of the IBM virtual storage environment.

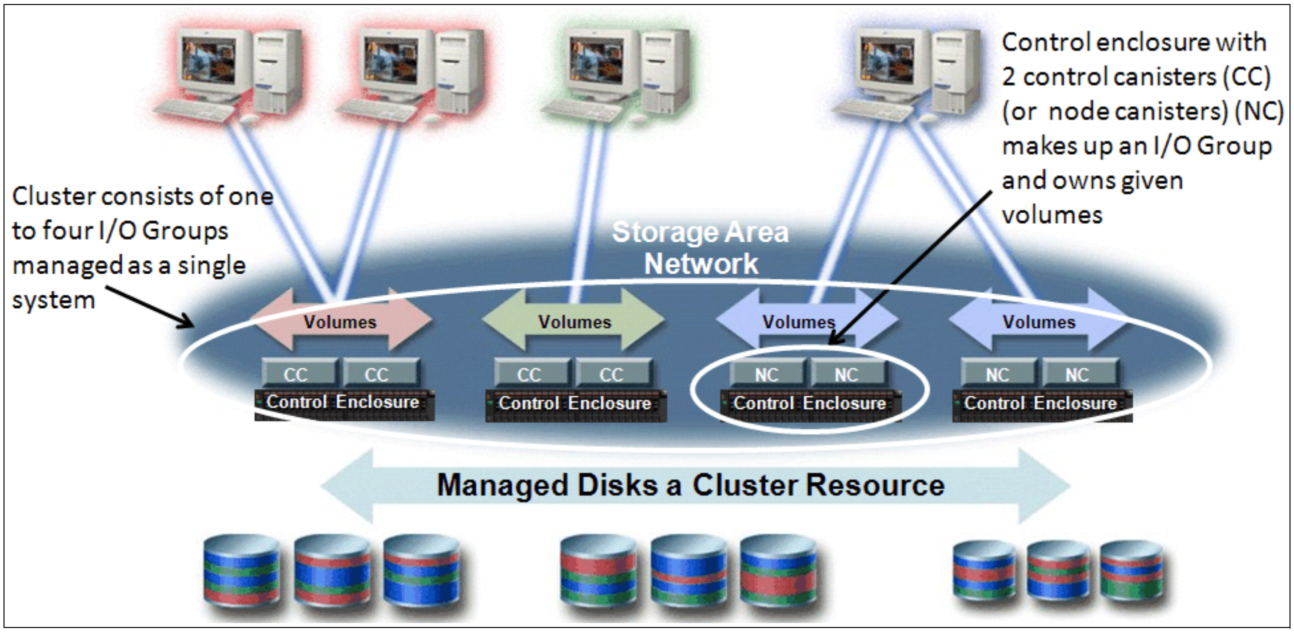


Figure 2-1 IBM virtual storage environment

2.2 IBM Flex System V7000 Storage Node terminology

The IBM Flex System V7000 Storage Node introduces some new terminology, which is defined below in Table 2-1. We also include the terms first introduced with the IBM SVC and the IBM Storwize V7000, and are important to understand the rest of the implementation procedures discussed in this publication.

Table 2-1 IBM Flex System V7000 Storage Node terminology

IBM Flex System V7000 Storage Node term	Definition
Chain	The SAS2 connections by which expansion enclosures are attached which provides redundant access to the drives that are inside the enclosures. Each IBM Flex System V7000 Storage Node control canister has one chain connection.
Clone	A copy of a volume on a server at a particular point in time. The contents of the copy can be customized while the contents of the original volume are preserved.
Control canister	A hardware unit that includes all the management and control hardware, fabric and service interfaces, and the SAS2 expansion port.
Control enclosure	A hardware unit chassis that inserts into the IBM Flex System that includes control canisters including backup batteries, and 24 drive slots. This is the initial building block for the IBM Flex System V7000 Storage Node.

IBM Flex System V7000 Storage Node term	Definition
Event	An occurrence that is significant to a task or system. Events can include completion or failure of an operation, a user action, or the change in the state of a process.
Expansion canister	A hardware unit that includes the serial-attached SCSI (SAS2) interface hardware that enables the control enclosure to use the drives of the expansion enclosure as well as other expansions to be daisy-chained on behind it.
Expansion enclosure	A hardware unit chassis that inserts into the IBM Flex System that includes expansion canisters, and 24 drive slots used for connecting additional internal capacity to the IBM Flex System V7000 Storage Control Enclosure.
External V7000 expansion	A 2076-212 or 2076-224 IBM Storwize V7000 expansion that is connected to the IBM Flex System V7000 Storage Control Enclosure via the SAS2 chain to provide additional storage capacity which resides outside of the IBM Flex System.
External Virtualized storage	Managed disks (MDisks) that are presented as logical drives by external storage systems that are to be attached to the IBM Flex System V7000 Storage Node for additional virtual capacity.
Host mapping	The process of controlling which hosts have access to specific volumes within a clustered system.
Internal storage	The storage that resides in the control and expansion enclosures or the IBM Storwize V7000 expansions connected through the SAS2 chain that make up the IBM Flex System V7000 Storage Node.
Lane	The name given for a single 6Gbps SAS2 PHY (channel). There are four lanes (PHY) that make up each SAS2 chain.
Managed disk (MDisk)	A component of a storage pool that is managed by a clustered system. An MDisk is either a RAID array created using the internal storage, or a Small Computer System Interface (SCSI) logical unit (LU) for external storage being virtualized. An MDisk is not visible to a host system on the storage area network.
PHY	A term used to define a single 6Gbps SAS lane. There are four PHYs in each SAS cable.
Quorum disk	A disk that contains a reserved area that is used exclusively for cluster management. The quorum disk is accessed when it is necessary to determine which half of the cluster continues to read and write data. Quorum disks can either be on an MDisks or internal drives.
Snapshot	An image backup type that consists of a point-in-time view of a volume.
Storage pool	A collection of storage capacity on mdisks that can be used to provide the capacity requirements for a volume.
Strand	The serial-attached SCSI (SAS) connectivity of a set of drives within multiple enclosures. The enclosures can be either the IBM Flex System V7000 Storage control enclosures or expansion enclosures, or the IBM Storwize V7000 expansion enclosures.
Thin provisioning or Thin provisioned	The ability to define a storage unit (full system, storage pool, or volume) with a logical capacity size that is larger than the physical capacity assigned to that storage unit.

IBM Flex System V7000 Storage Node term	Definition
Volume	As used with the IBM Flex System V7000 Storage Node environment, this is the virtually defined device created for use by the host or the IBM Flex System V7000 Storage Node cluster to store IO data.

2.3 IBM Flex System V7000 Storage Node

The IBM Flex System V7000 Storage Node is based on two enclosure types; IBM Flex System V7000 Control Enclosure and the IBM Flex System V7000 Expansion Enclosure. Both of these enclosures reside in a newly designed common chassis that fits into the IBM Flex System Enterprise chassis. This new enclosure requires a space equal to two high and double wide slots be available for each internal enclosure. Figure 2-2 shows an IBM Flex System V7000 Storage Node that has been built with a control enclosure and one expansion enclosure.

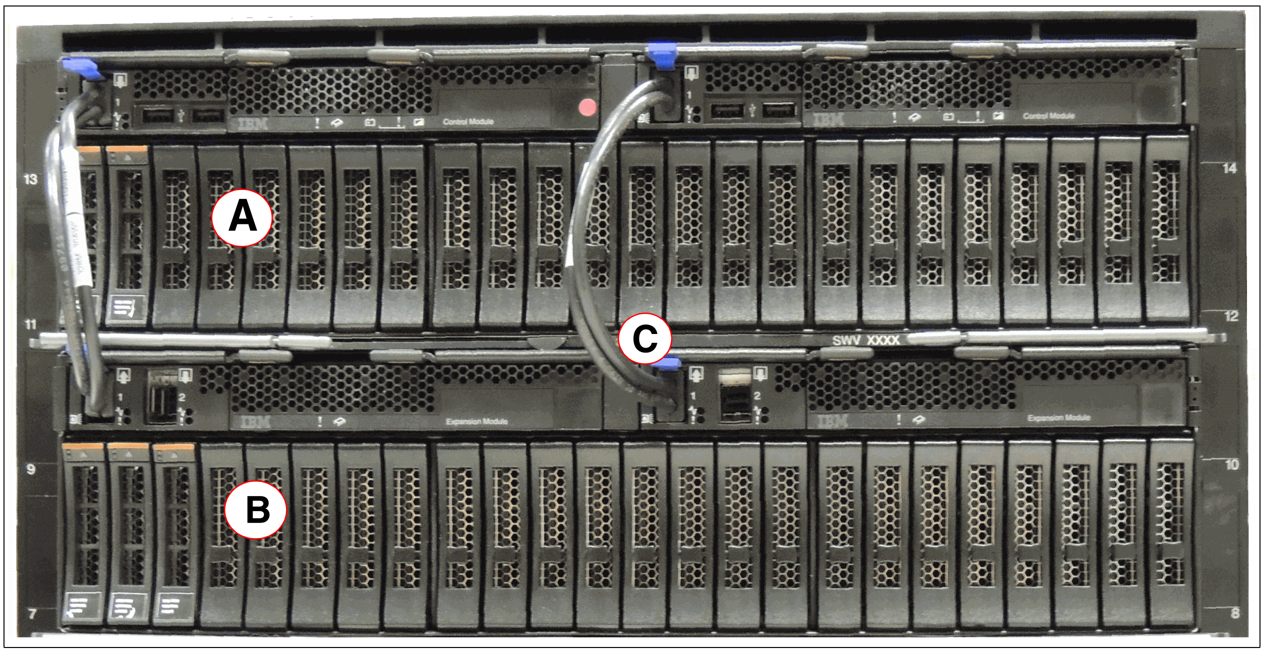


Figure 2-2 IBM Flex System V7000 Storage Node

Both the control and expansion enclosures connect to the Flex System Enterprise chassis through the midplane interconnect for their power and internal control connections. The control enclosure (A) also connects to the IBM Flex System IO modules and switches for host IO and replication features through this midplane. The control enclosure also houses a pair of redundant control canisters along with their cache batteries for backup.

The expansion enclosure (B) uses the Serial Attached SCSI (SAS2) 6 Gbps chain connection on the front of the control and expansion canisters (C) for connecting the chain together for drive IO and expansion control operations.The expansion enclosure houses a pair of expansion canisters instead of the control canisters through which it connects and manages the SAS chain connections to its disk drives. It also has a second SAS2 port through which it provides a connection for continuing the chain to additional expansions behind it.

An IBM Flex System V7000 Storage Node is mainly intended to be a scalable, internal storage system, to support the internal compute nodes of the IBM Flex System. When needed, it can be expanded in its capacity by attaching external IBM Storwize V7000 expansion enclosures to its SAS2 chain. Both the 2076-212 and the 2076-224 model of the Storwize V7000 expansions are supported.

The control enclosure can support a combination of up to nine expansion enclosures using a combination of internal IBM Flex System V7000 Expansion Enclosures (max 2)and external IBM Storwize V7000 Expansion Enclosures (max 9) connected through the SAS2 connection on each of the control canisters. With the additional expansions you are capable of up to 240 disks drives; or a maximum capacity of 348TB per control enclosure.

2.3.1 IBM Flex System V7000 Storage Node capabilities

For a complete and updated list of IBM Flex System V7000 Storage Node configuration limits and restrictions, see the following address:

http://www-01.ibm.com/support/docview.wss?uid=ssg1S1004114

2.3.2 IBM Flex System V7000 Storage Node functions

The following functions are available with IBM Flex System V7000 Storage Node:

- Thin provisioning (No license required)

  Traditional fully allocated volumes allocate real physical disk capacity for an entire volume even if that capacity is never used. Thin-provisioned volumes allocate real physical disk capacity only when data is written to the logical volume.

- Volume mirroring (No license required)

  Provides a single volume image to the attached host systems while maintaining pointers to two copies of data in separate storage pools. Copies can be on separate disk storage systems that are being virtualized. If one copy is failing, IBM Flex System V7000 Storage Node provides continuous data access by redirecting I/O to the remaining copy. When the copy becomes available, automatic re-synchronization occurs.

- FlashCopy (included with the base IBM Flex System V7000 Storage Node license)

  Provides a volume level point-in-time copy function for any storage being virtualized by IBM Flex System V7000 Storage Node. This function is designed to create copies for backup, parallel processing, testing, and development, and have the copies available almost immediately.

  IBM Flex System V7000 Storage Node includes the following FlashCopy functions:

  – Full / Incremental copy

    This function copies only the changes from either the source or target data since the last FlashCopy operation and is designed to enable completion of point-in-time online backups much more quickly than using traditional FlashCopy.

  – Multitarget FlashCopy

    IBM Flex System V7000 Storage Node supports copying of up to 256 target volumes from a single source volume. Each copy is managed by a unique mapping and, in general, each mapping acts independently and is not affected by other mappings sharing the source volume.

  – Cascaded FlashCopy

This function is used to create copies of copies and supports full, incremental, or nocopy operations.

- Reverse FlashCopy

  This function allows data from an earlier point-in-time copy to be restored with minimal disruption to the host.

- FlashCopy nocopy with thin provisioning

  This function provides a combination of using thin-provisioned volumes and FlashCopy together to help reduce disk space requirements when making copies. There are two variations of this option:

  - Space-efficient source and target with background copy

    Copies only the allocated space.

  - Space-efficient target with no background copy

    Copies only the space used for changes between the source and target and is generally referred to as “snapshots”.

  This function may be used with multi-target, cascaded, and incremental FlashCopy.

- Consistency groups

  Consistency groups address the issue where application data is on multiple volumes. By placing the FlashCopy relationships into a consistency group, commands can be issued against all of the volumes in the group. This action enables a consistent point-in-time copy of all of the data, even though it might be on a physically separate volume.

  FlashCopy mappings can be members of a consistency group, or they can be operated in a stand-alone manner, that is, not as part of a consistency group. FlashCopy commands can be issued to a FlashCopy consistency group, which affects all FlashCopy mappings in the consistency group, or to a single FlashCopy mapping if it is not part of a defined FlashCopy consistency group.

- Remote Copy feature

  Remote Copy is an optional licensed feature that is based on the number of enclosures that are being used at the smallest configuration location. See 2.6.3, “Remote Copy licensing” on page 74 for licensing details. Remote Copy provides for the capability to perform either Metro mirror or Global Mirror operations.

  - Metro Mirror

    Provides a synchronous remote mirroring function up to approximately 300 km between sites. As the host I/O only completes after the data is cached at both locations, performance requirements might limit the practical distance. Metro Mirror is designed to provide fully synchronized copies at both sites with zero data loss after the initial copy is completed.

    Metro Mirror can operate between multiple IBM Flex System V7000 Storage Node systems and is only supported on either FC or FCoE host interfaces.

  - Global Mirror

    Provides long distance asynchronous remote mirroring function up to approximately 8,000 km between sites. With Global Mirror, the host I/O completes locally and the changed data is sent to the remote site later. This function is designed to maintain a consistent recoverable copy of data at the remote site, which lags behind the local site.

    Global Mirror can operate between multiple IBM Flex System V7000 Storage Node systems and is only supported on FC host interfaces.

- Data Migration (no charge for temporary usage)

  IBM Flex System V7000 Storage Node provides a data migration function that can be used to import external storage systems into the IBM Flex System V7000 Storage Node system.

  It allows you to:

  – Move volumes non disruptively onto a newly installed storage system

  – Move volumes to rebalance a changed workload

  – Migrate data from other back-end storage to IBM Flex System V7000 Storage Node managed storage

- IBM System Storage Easy Tier (no charge)

  Provides a mechanism to seamlessly migrate hot spots to the most appropriate tier within the IBM Flex System V7000 Storage Node solution. This migration could be to internal drives within IBM Flex System V7000 Storage Node or to external storage systems that are virtualized by IBM Flex System V7000 Storage Node.

- Real-time Compression (RTC)

  Provides for data compression using the IBM Random-Access Compression Engine (RACE) which can be performed on a per volume basis in real time on active primary workloads. RTC can provide as much as a 50% compression rate for data that is not already compressed. This can help with reducing the amount of capacity needed for storage which can help with delaying further growth purchases. RTC supports all storage that is attached to the IBM Flex System V7000 Storage Node whether internal, external, or external virtualized storage.

  A compression evaluation tool called the IBM Comprestimator can be used to determine the value of using compression on a specific workload for your environment.

- External Storage Virtualization (Licensed per enclosure of the external storage subsystem)

  With this feature an external storage subsystem can be attached through the fiber channel or by FCoE to the IBM Flex System V7000 Storage Node. These devices cannot be presented through an iSCSI connection. The devices presented are treated as mdisks and can be mapped to storage pools for volume creation and management. Once the storage from the external system is integrated into the IBM Flex System V7000 Storage Node and added to a storage pool it is available to be virtualized and used by any of the features and functions of the IBM Flex System V7000 Storage Node.

  External virtualization: External Virtualization is supported only on FC and FCoE but not iSCSI.

2.3.3 IBM Flex System V7000 Storage Node licensing

IBM Flex System V7000 Storage Node might require the following licenses:

- Enclosure
- External Virtualization
- Remote Copy (Advanced Copy Services: Metro Mirror / Global Mirror)
- Real-time Compression (RTC)

Table 2-2 on page 47 gives a summary of all the licenses that might be required.

Table 2-2 Licenses that might be required

License type	Unit	License name	License required?
Enclosure	Base+expansion Physical Enclosure Number	IBM Flex System V7000 Base Software	Yes, software license per enclosure.
External Virtualization	Physical Enclosure Number Of External Storage	IBM Flex System V7000 External Virtualization Software	Optional addon feature Yes, software license per external storage enclosure.
Remote Copy	Physical Enclosure Number	IBM Flex System V7000 Remote Mirroring Software	Optional addon feature Yes, software license per enclosure.
Real-time Compression (RTC)	Physical Enclosure Number	IBM Flex System V7000 Real-time Compression Software	Optional addon feature Yes, software license per enclosure.
FlashCopy	N/A	N/A	No
Volume Mirroring	N/A	N/A	No
Thin Provisioning	N/A	N/A	No
Volume Migration	N/A	N/A	No
Easy Tier	N/A	N/A	No

Migration: If the IBM Flex System V7000 Storage Node is used as a general migration tool, then the appropriate External Virtualization licenses must be ordered. The only exception is if you want to migrate existing data from external storage to IBM Flex System V7000 Storage Node permanently managed storage; you can temporarily configure your External Storage license for use up to 45 days. For a more-than-45-day migration requirement from external storage to the IBM Flex System V7000 Storage Node the appropriate External Virtualization license must be ordered.

2.4 IBM Flex System V7000 Storage Node hardware

The IBM Flex System V7000 Storage Node solution is a modular storage system that is built to interface and reside within the IBM Flex System Enterprise Chassis. When sold in the IBM PureFlex system it will come with a four port FC host interface card (HIC) preinstalled in it and configured to the FC switch(s) in the Flex Enterprise Chassis. For the “Build to Order” (BTO) solutions you can select the configuration of the interface capability you require for your specific environment.

Host connectivity to compute nodes is provided through optional Flex System V7000 control enclosure network cards that connect to the Flex System Enterprise Chassis midplane and its switch modules. Available options are:

- 10Gb Converged Network Adapter (CNA) 2 Port Daughter Card for FCoE and iSCSI fabric connections
- 8Gb Fibre Channel (FC) 4 Port Daughter Card for Fibre Channel fabric connections.

The Flex System V7000 Storage Node has two available slots in each control canister for populating with host network adapters. The adapters in each of the node canisters must be of the same type in each of the slots; so the adapters are installed in pairs (one adapter per node canister) and the following adapter configurations are supported for the entire Flex System V7000 Storage Node:

- Two or four 10Gb CNA network adapters
- Two 4 port 8Gb FC network adapters
- Two 10Gb CNA network adapters and two 4 port 8Gb FC network adapters

The 2 port 10 Gb Ethernet network adapters (up to two per canister) are used for iSCSI host attachment and/or FCoE attachment.

The configuration of the host attachments on one control canister must match the configuration of the second.

There is a 6 Gbps SAS port on the front of the canister for connecting optional expansions enclosures.

2.4.1 Control canister

The control canister is responsible for the management of all the virtualization, RAID functions, advanced features and functions of the IBM Flex System V7000 Storage Node and all command and IO to its internal drive slots and the expansions that it is connected to. The control canister is a Customer Replaceable Unit (CRU). Figure 2-3 is a picture of the control canister with its cover removed.

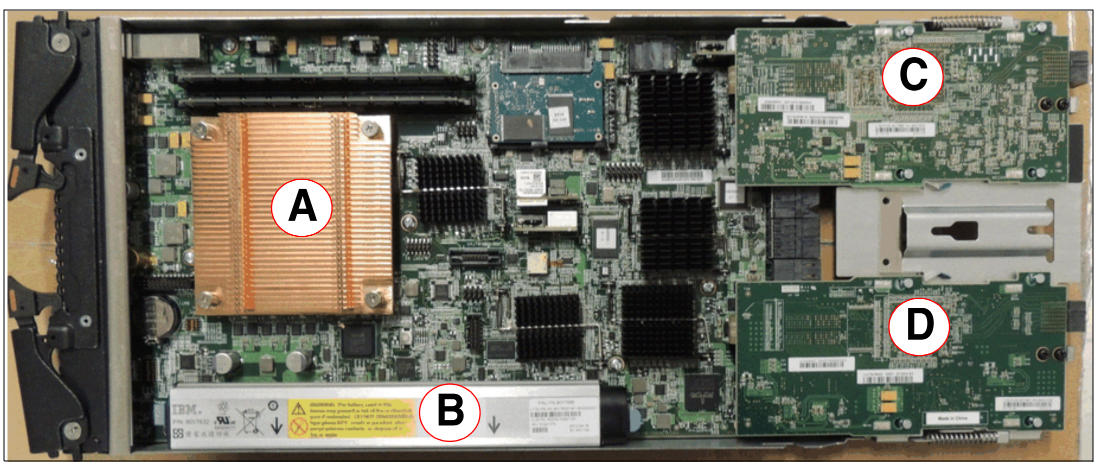


Figure 2-3 Components and board layout of the control canister

In Figure 2-3, the processor (A) is a quad core Jasper. There are also two DIMMs which make up the cache memory and the battery backup unit (B) which beside the control canister is also a Customer Replaceable Unit (CRU).

As we discussed above, the two host network adapter card locations (C and D) can also be seen. These provide the connections to the Flex System Enterprise Chassis through the midplane and the switch modules. It is important to remember that both control canisters must be populated with the same type of host network adapter cards in each of these locations.

Attention: The first network adapter (slot1) location can only be populated by a 2 port 10Gbps ethernet CNA network adapter; the second location can be populated by either a 2 port 10Gbps ethernet CNA network adapter or a 4 port 8Gbps fiber channel network adapter.

Figure 2-4 is a logical block diagram of the components and flow of the control canister

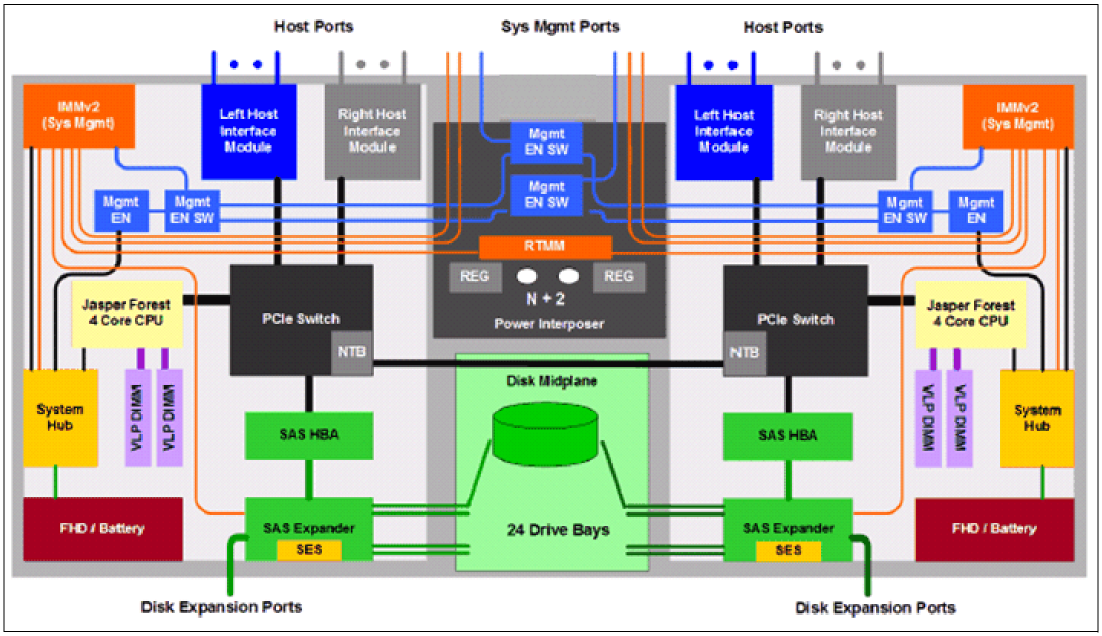


Figure 2-4 Logical block diagram for control canister

As shown in Figure 2-5 on page 50, the control canister has only one SAS2 ports per canister for the connection of expansion enclosures to add capacity. There are also two USB connections available for support use to perform maintenance actions as required.

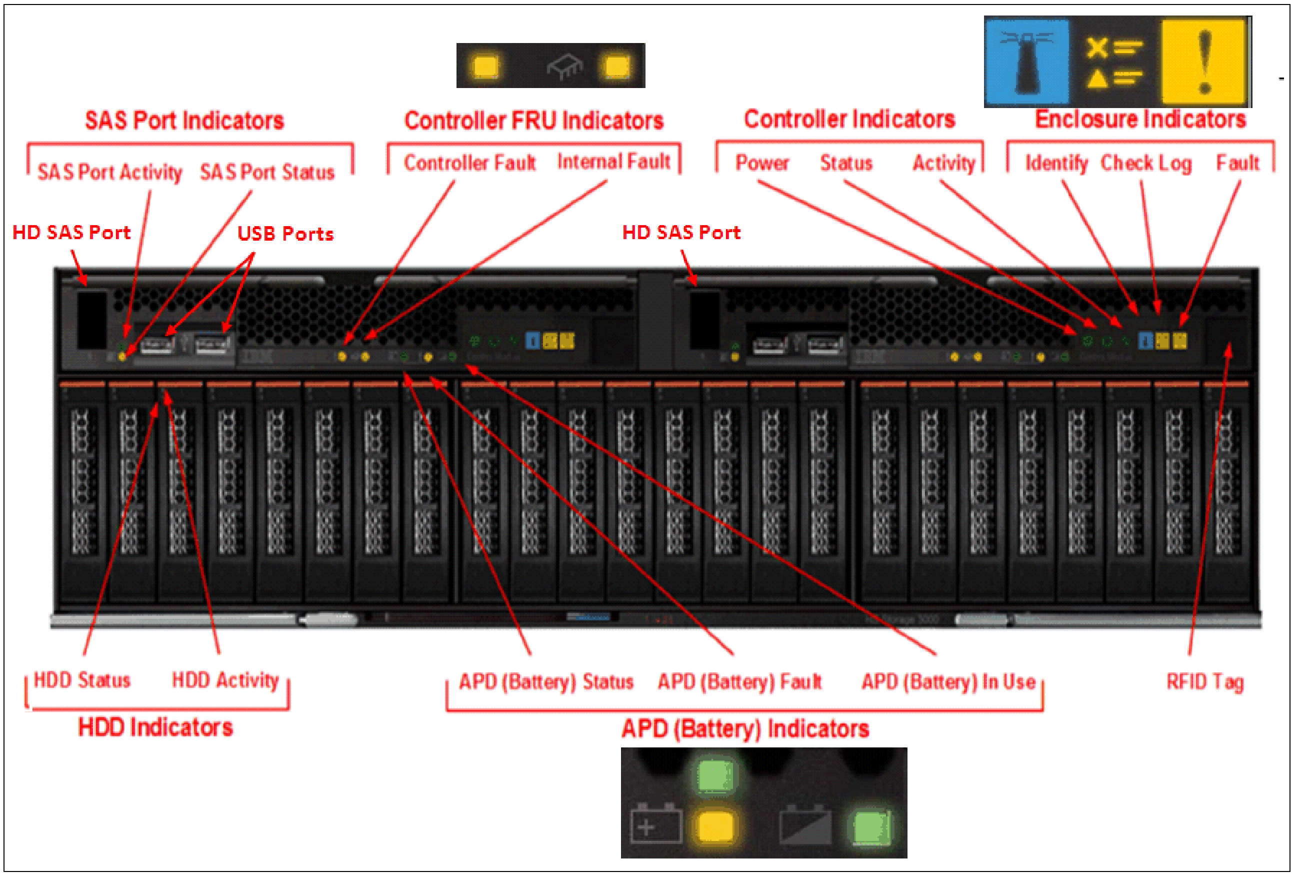


Figure 2-5 Control enclosure with connection ports and indicators

In Table 2-3, the LED indicators are defined for the control enclosure; some of these are also used on the expansion enclosure as well.

Table 2-3 Control enclosure LED description

LED Group	LED Name	Meaning
Enclosure Indicator The Enclosure Indicator LEDs on the righthand control canister are used as the primary enclosure LEDs. The lefthand control canister LEDs are used when the righthand is not available.	Check Log	Is a software fault telling you to check the errorlog for details.
	Identify	Used for identifying the selected enclosure.
	Fault	Indicates a hardware fault has occurred.

<table>
<tr><th>LED Group</th><th>LED Name</th><th>Meaning</th></tr>
<tr><td rowspan="3">Controller Indicators</td><td>Power</td><td>► Slow Blink - Power available but Processor shutdown.
► Fast blink - doing POST.
► On solid - Powered up.</td></tr>
<tr><td>Status</td><td>► Off - not operational
► Solid - in a cluster
► Slow blinking - in cluster or service state
► Fast blink - is upgrading</td></tr>
<tr><td>Activity</td><td>The Activity LED blinks to show there is IO activity</td></tr>
<tr><td rowspan="2">Controller FRU

Control canister hardware fault</td><td>Control</td><td>Hardware problem with the control canister.</td></tr>
<tr><td>Internal FRU</td><td>Problem with an internal FRU like the network adapters.</td></tr>
<tr><td rowspan="3">Battery Indicators</td><td>In Use</td><td>Fast blink- System is shutting down on battery power.</td></tr>
<tr><td>Status</td><td>► Slow blink - Charging, but has enough power to boot the canister.
► Fast blink - Charging, and does not have enough power to boot the canister (Error 656)
► On solid - Fully charged.</td></tr>
<tr><td>Fault</td><td>On - Hardware fault detected</td></tr>
</table>

2.4.2 Expansion canister

The expansion canister connects the expansion disks to the control canister using the SAS2 6 Gbps chain interface. This module also enables the daisy-chaining of additional expansions to be connected behind it to further expand the capacity of the system's storage. The expansion canister is also a Customer Replaceable Unit (CRU). Figure 2-6 on page 52 is a picture of the expansion canister with its cover removed.

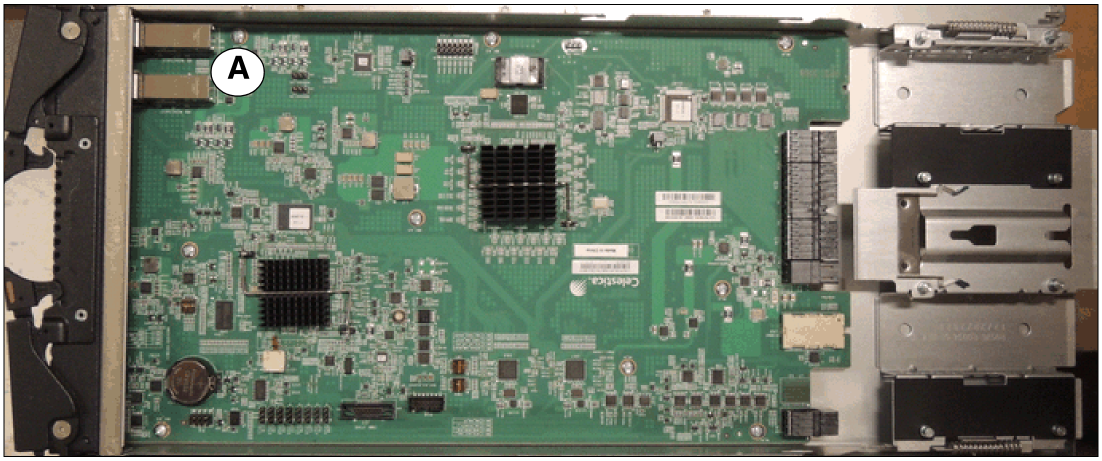


Figure 2-6 Components and board layout of the expansion canister

As shown in Figure 2-6, the expansion canister does not contain a battery backup unit like the control canister does. It does have an additional SAS2 connection (A) on it to allow for the continuation of the chain to additional expansions. Figure 2-7 shows the SAS connection ports and the indicators on the front of the expansion enclosures.

Confirm that their is a Status indicator on the expansion enclosures.

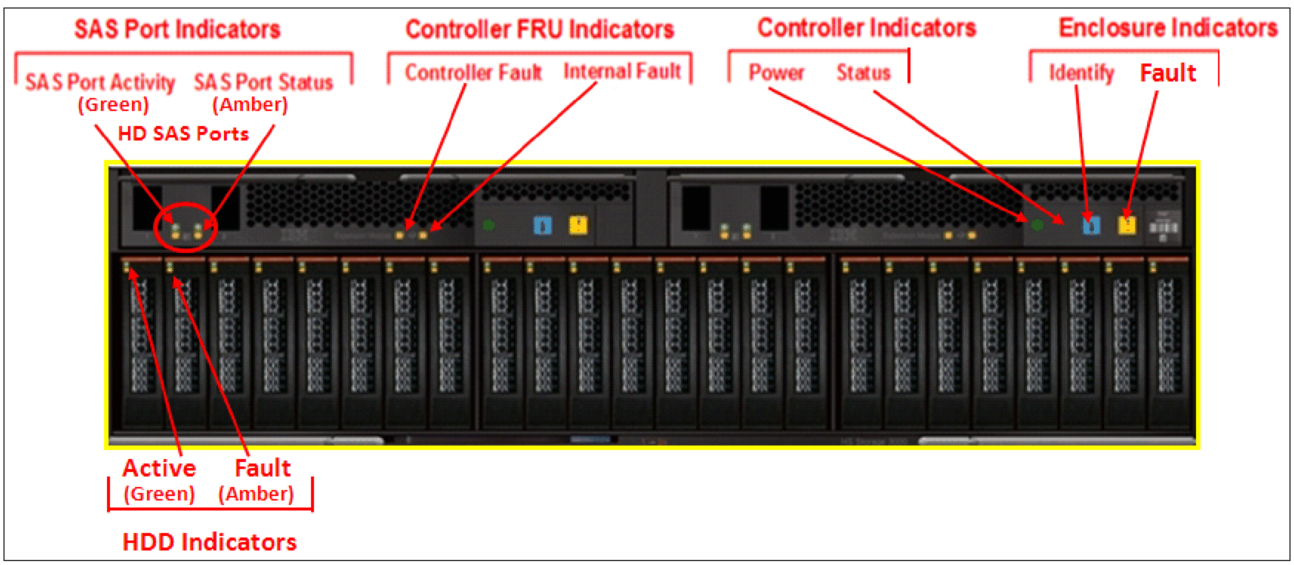


Figure 2-7 Expansion enclosure with connection ports and indicators

As the indicators on the expansion enclosure are a subset of the ones that are available on the control enclosure, Table 2-3 on page 50 provides the details of their definitions.

Figure 2-8 on page 53 is a logical block diagram of the expansion canister.

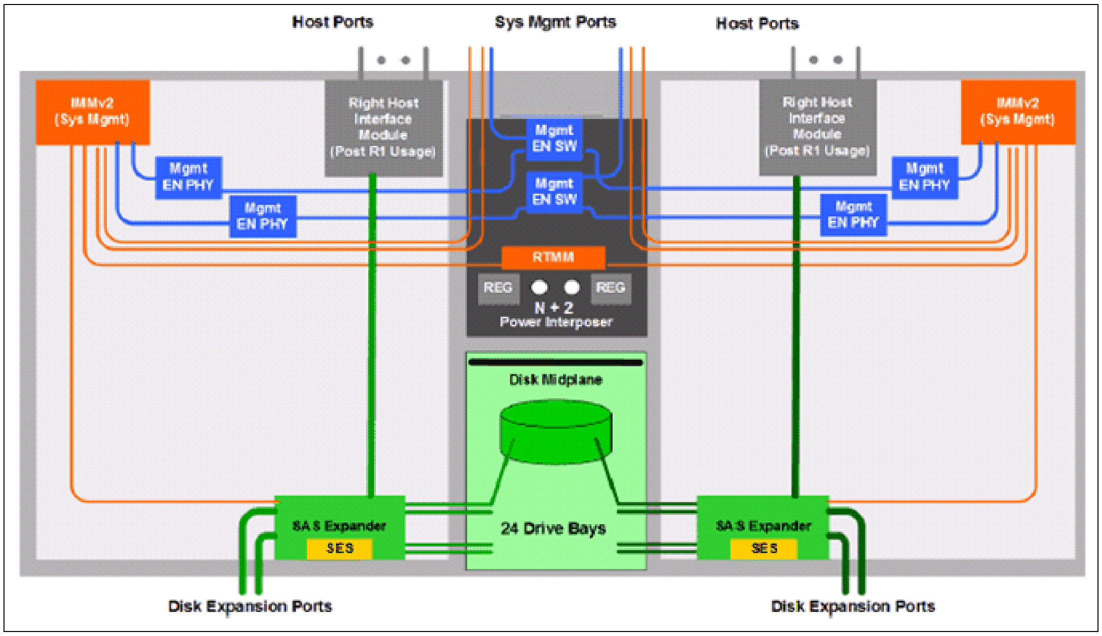


Figure 2-8 Logical block diagram for control canister

2.4.3 Supported disk drives

Both the IBM Flex System V7000 Control Enclosure and the IBM Flex System V7000 Expansion Enclosure support up to 24 2.5 inch disk drives in their enclosures. Table 2-4 shows all the possible drive types that can be used in the internal enclosure drive slots at the time of writing.

Table 2-4 IBM Flex System V7000 Storage Node internal supported drives

Drive Capacity	Drive speed	Drive physical size
146 GB	15 KRPM SAS	2.5 inch
300 GB	15 KRPM SAS	2.5 inch
300 GB	10 KRPM SAS	2.5 inch
600 GB	10 KRPM SAS	2.5 inch
900 GB	10 KRPM SAS	2.5 inch
500 GB	7500 RPM NL SAS	2.5 inch
1 TB	7500 RPM NL SAS	2.5 inch
200 GB	SSD SAS	2.5 inch
400 GB	SSD SAS	2.5 inch

The disk drives connect to the Flex System Enterprise chassis through the midplane interconnect for their power. Also in the control enclosure, the midplane interconnect is used for the internal control and IO paths. The expansion enclosures use Serial Attached SCSI (SAS) connections on the front of the control and expansion canisters for disk IO and control commands.

2.4.4 IBM Storwize V7000 expansion enclosure

The IBM Storwize V7000 expansion enclosure can be optionally attached to the IBM Flex System V7000 Storage Node for added capacity beyond that of the internal enclosures. These expansion enclosures contain two IBM Storwize V7000 *expansion canisters*, disk drives, and two power supplies. There are two models of the expansion enclosures with the 2076-212 providing 12 disk slots of the 3.5 inch form factor, and the 2076-224 providing 24 disk slots of the 2.5 inch drives form factor.

Figure 2-9 shows the components of the expansion enclosure.

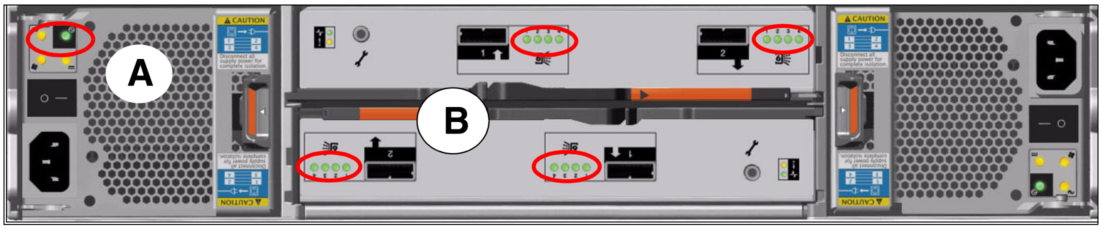


Figure 2-9 Component side view Expansion enclosure

The expansion enclosure power supplies have a single power lead connector on the power supply unit. The PSU has an IEC C14 socket and the mains connection cable has a C13 plug. As shown in Figure 2-9, the PSU has one green status LED indicator (A) to show it is powered on and working properly.

Each expansion canister provides two SAS interfaces that are used to connect to either the IBM Flex System V7000 Storage Node control or expansion enclosures or to a preceding IBM Storwize V7000 expansion enclosure, as well as to connect any additional IBM Storwize V7000 expansion enclosure behind it. The ports are numbered 1 on the left and 2 on the right. SAS port 1 is the IN port and SAS port 2 is the OUT port. There is also a symbol printed at the ports to identify whether it is an IN or an OUT bound port.

Use of the SAS connector 1 is mandatory when installed, as the expansion enclosure must be attached to either a control enclosure or another expansion enclosure. SAS connector 2 is optional, as it is used to attach to the next additional expansion enclosure in the chain.

Each port connects four SAS physical links (PHYs) to the drives. As shown in Figure 2-9 there is an LED associated with each PHY in each port (eight LEDs in total per canister). The LEDs are green and grouped together next to the ports (B); for each port they are numbered 1 through 4. These LED's indicate when there is activity on the PHY.

Figure 2-10 shows the front view of the 2076-212 enclosure.

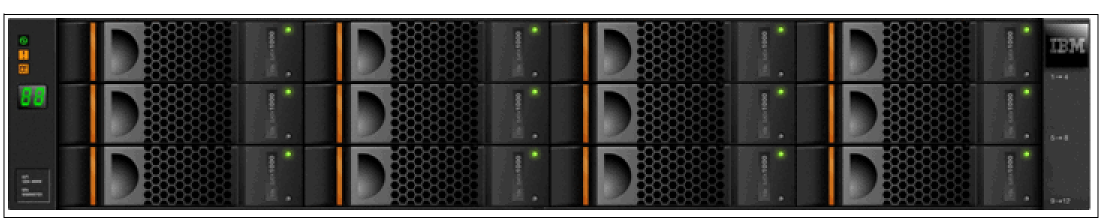


Figure 2-10 IBM Storwize V7000 front view for 2076-212enclosure

The drives are positioned in four columns of three horizontally mounted drive slots in the expansion enclosure. The drive slots are numbered 1 - 12, starting at the upper left and going from left to right, top to bottom.

Figure 2-11 shows the front view of the 2076-224, enclosure.

Figure 2-11 IBM Storwize V7000 front view for 2076-224, enclosure

The drives are positioned in one row of 24 vertically mounted row in the drive assemblies. The drive slots are numbered 1 - 24, starting from the left. (There is a vertical center drive bay molding between slots 12 and 13).

Though the IBM Storwize V7000 2076-224 enclosure is a 2.5 inch, 24 drive slot chassis; the drives of this subsystem are not interchangeable with the IBM Flex System V7000 Storage Node drives. Both drives use a different carrier and contain a different product identifier in their code.

Therefore, the IBM Storwize V7000 expansions support their own drive types in their enclosures and should be created in their own configurations and storage pools when used with the IBM Flex System V7000 Storage Node.

IBM Flex System V7000 Storage Node enclosures currently support SSD, SAS, and Nearline SAS drive types. Each SAS drive has two ports (two PHYs) and I/O can be issued down both paths simultaneously.

For a list of supported drives for the IBM Storwize V7000 see the InfoCenter website for details at:

http://pic.dhe.ibm.com/infocenter/storwize/ic/index.jsp

2.4.5 SAS cabling requirements

The IBM Flex System V7000 Storage Node uses new smaller SAS2 cable connectors. These connectors are based on the high density (HD) mini SAS connectors. For the connections between the IBM Flex System V7000 Control enclosure and the IBM Flex System V7000 Expansion enclosure in the chassis, you have to use these short cables (as shown in Figure 2-2 on page 43) that can be ordered with your storage.

With the addition of external IBM Storwize V7000 expansion there is an adapter which changes the HD mini SAS to the mini SAS connection that is on the expansion. Figure 2-12 on page 56 shows the cabling scheme and the differences in the cable connections of the two expansions.

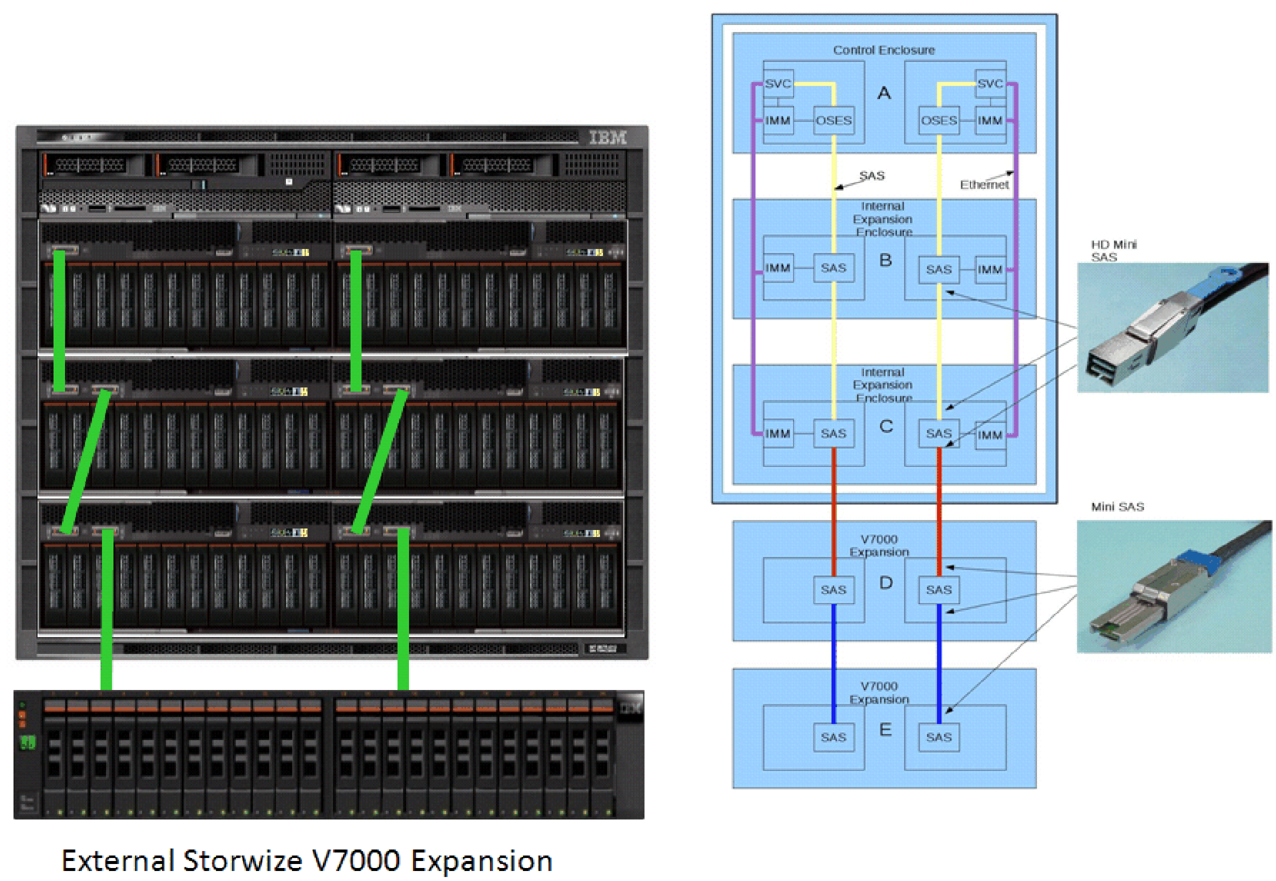


Figure 2-12 IBM Flex System V7000 Storage Node cabling internal SAS and external

It should be noted that the IBM Flex System V7000 Storage Node cable directly from top to bottom down through its expansions including any additional external expansions.

2.5 IBM Flex System V7000 Storage Node components

The IBM Flex System V7000 Storage Node is an integrated entry / midrange virtualization RAID storage to support and interface with the IBM Flex System. It brings with it the following benefits:

- Single point of management for both server, network and internal storage and external storage virtualized by the integrated IBM Flex System V7000 Storage Node.
- Improved access to critical data by keeping data close to the servers by minimizing switch and cable hops.
- Provides virtual servers fast shared storage access to better enable dynamic workload assignments.
- Moved the control of the virtual server storage over to the server administrators rather than in the hands of the SAN administrators.
- Leverage the IBM Flex System Enterprise Chassis for lower solution costs eliminating the need of external switches, cables, SFPs, fans, power supplies, for production environments.

Note: Even though there is no external switches required there is a requirement for the appropriate internal switch IO modules to be available for the desired host interfacing to work.

The IBM Flex System V7000 Storage Node consists of a set of drive enclosures. Control enclosures contain disk drives and two control canisters and form an I/O group for its management of additional internal and external storage. Expansion enclosures contain drives and are attached to control enclosures.

The simplest use of the IBM Flex System V7000 Storage Node is as a traditional RAID subsystem. The internal drives are configured into RAID arrays and virtual disks created from those arrays.

The IBM Flex System V7000 Storage Node can also be used to virtualize other storage controllers, and an example of this is described in Chapter 7, “Migration wizard” on page 265.

The IBM Flex System V7000 Storage Node supports regular and solid-state drives and can use IBM System Storage Easy Tier to automatically place volume hot spots on better-performing storage.

2.5.1 Hosts

A host system is an IBM Flex System compute node that is connected to IBM Flex System V7000 Storage Node through a fiber channel connection or through an ethernet connection using either iSCSI or Fiber Channel over Ethernet (FCoE) connection through the switch modules of the IBM Flex System. At the time of writing, attachment of external hosts to the IBM Flex System V7000 Storage Node is not supported.

Hosts are defined to IBM Flex System V7000 Storage Node by identifying their worldwide port names (WWPNs) for Fibre Channel hosts. For iSCSI hosts, they are identified by using their iSCSI names. The iSCSI names can either be iSCSI qualified names (IQNs) or extended unique identifiers (EUIs).

2.5.2 Control canisters

IBM Flex System V7000 Storage Node can have two to eight hardware components called control canisters that provide the virtualization of internal and external volumes, and cache and copy services (Remote Copy) functions. Within an IBM Flex System V7000 Storage Node a pair of the control canisters are housed within a control enclosure. A clustered system consists of a one to four control enclosures.

One of the control canisters within the system is known as the configuration node and it is the canister that manages configuration activity for the clustered system. If this canister fails, the system nominates another canister to become the configuration node. During initial setup the system will automatically select a canister for this role from the first control enclosure pair.

2.5.3 I/O groups

Within IBM Flex System V7000 Storage Node there can be one to four control enclosures which also are defined as *I/O groups*. The IBM Flex System V7000 Storage Node supports four control enclosures in the clustered system, which provides four I/O groups.

When a host server performs I/O to one of its volumes, all the I/Os for a specific volume are directed to the owning I/O group. Also, under normal conditions, the I/Os for that specific volume are always processed by the same canister within the I/O group.

Both canisters of the I/O group act as preferred nodes for their own specific subset of the total number of volumes that the I/O group presents to the host servers (a maximum of 2048 volumes). However, both canisters also act as the failover canister for its partner within the I/O group, so a canister can take over the I/O workload from its partner, if required.

In a IBM Flex System V7000 Storage Node environment, using active-active architecture, the I/O handling for a volume can be managed by both canisters of the I/O group. Therefore, it is mandatory for servers that are connected through Fibre Channel connectors to use multipath device drivers to be able to handle this capability.

The I/O groups are connected to the SAN so that all application servers accessing volumes from the I/O group have access to them. Up to 512 host server objects can be defined in two I/O groups.

Important: The active / active architecture provides ability to process I/Os for both control canisters and allows the application to continue running smoothly, even if the server has only one access route or path to the storage. This type of architecture eliminates the path / LUN thrashing that can exist with an active / passive architecture.

2.5.4 Clustered system

A clustered system consists of one to four control enclosures. All configuration, monitoring, and service tasks are performed at the system level and the configuration settings are replicated across all control canisters in the clustered system. To facilitate these tasks, one or two management IP addresses are set for the system.

There is a process provided to back up the system configuration data on to disk so that the clustered system can be restored in the event of a disaster. This method does not back up application data, only the IBM Flex System V7000 Storage Node configuration information.

System configuration backup: After backing up the system configuration, save the backup data on your hard disk (or at the least outside of the SAN). When you are unable to access the IBM Flex System V7000 Storage Node, you do not have access to the backup data if it is on the SAN.

For the purposes of remote data mirroring, two or more clustered systems (IBM Flex System V7000 Storage Nodes, Storwize V7000, or SAN Volume Controller systems starting from Version 6.4) must form a partnership before creating relationships between the mirrored volumes.

Important: With IBM Flex System V7000 Storage Node 6.4, the **`layer`** parameter can only be changed by running **`chsystem`** using the CLI. The default is the `storage` layer, and you must change it to `replication` if you need to set up a copy services relationship between the IBM Flex System V7000 Storage Node with either the IBM Storwize V7000 or SAN Volume Controller.

As mentioned earlier, one canister is designated as the configuration node and it is the canister that activates the system IP address. If the configuration node fails, the system

chooses a new canister as the configuration node and the new canister takes over the system IP addresses.

The system can be configured using either the IBM Flex System V7000 Storage Node management software the command-line interface (CLI),

2.5.5 RAID

The IBM Flex System V7000 Storage Node setup contains a number of internal disk drive objects known as candidate drives, but these drives cannot be directly added to storage pools. The drives need to be included in a Redundant Array of Independent Disks (RAID) grouping used for performance and to provide protection against the failure of individual drives.

These drives are referred to as members of the array. Each array has a RAID level. Different RAID levels provide different degrees of redundancy and performance, and have different restrictions regarding the number of members in the array.

IBM Flex System V7000 Storage Node supports hot spare drives. When an array member drive fails, the system automatically replaces the failed member with a hot spare drive and rebuilds the array to restore its redundancy (the exception being RAID 0). Candidate and spare drives can be manually exchanged with array members.

Each array has a set of goals that describe the wanted location and performance of each array. A sequence of drive failures and hot spare takeovers can leave an array unbalanced, that is, with members that do not match these goals. The system automatically rebalances such arrays when the appropriate drives are available.

The available RAID levels are:

- RAID 0 (striping, no redundancy)
- RAID 1 (mirroring between two drives)
- RAID 5 (striping, can survive one drive fault)
- RAID 6 (striping, can survive two drive faults)
- RAID 10 (RAID 0 on top of RAID 1)

RAID 0 arrays stripe data across the drives. The system supports RAID 0 arrays with just one member, which is similar to traditional JBOD attach. RAID 0 arrays have no redundancy, so they do not support hot spare takeover or immediate exchange. A RAID 0 array can be formed by one to eight drives.

RAID 1 arrays stripe data over mirrored pairs of drives. A RAID 1 array mirrored pair is rebuilt independently. A RAID 1 array can be formed by two drives only.

RAID 5 arrays stripe data over the member drives with one parity strip on every stripe. RAID 5 arrays have single redundancy. The parity algorithm means that an array can tolerate no more than one member drive failure. A RAID 5 array can be formed by 3 to 16 drives.

RAID 6 arrays stripe data over the member drives with two parity stripes (known as the P-parity and the Q-parity) on every stripe. The two parity strips are calculated using different algorithms, which give the array double redundancy. A RAID 6 array can be formed by 5 to 16 drives.

RAID 10 arrays have single redundancy. Although they can tolerate one failure from every mirrored pair, they cannot tolerate two-disk failures. One member out of every pair can be rebuilding or missing at the same time. A RAID 10 array can be formed by 2 to 16 drives.

2.5.6 Managed disks

A managed disk (MDisk) refers to the unit of storage that IBM Flex System V7000 Storage Node virtualizes. This unit could be a logical volume from an external storage array presented to the IBM Flex System V7000 Storage Node or a RAID array created on internal drives, or an external Storwize V7000 expansion that is managed by the IBM Flex System V7000 Storage Node. The IBM Flex System V7000 Storage Node can allocate these MDisks into various storage pools for different usage or configuration needs.

An MDisk should not be visible to a host system on the storage area network, as it should only be zoned to the IBM Flex System V7000 Storage Node system.

An MDisk has four possible modes:

- Array

  Array mode MDisks are constructed from drives using the RAID function. Array MDisks are always associated with storage pools.

- Unmanaged

  Unmanaged MDisks are not being used by the system. This situation might occur when an MDisk is first imported into the system, for example.

- Managed

  Managed MDisks are assigned to a storage pool and provide extents so that volumes can use it.

- Image

  Image MDisks are assigned directly to a volume with a one-to-one mapping of extents between the MDisk and the volume. This situation is normally used when importing logical volumes into the clustered system that already have data on them, which ensures that the data is preserved as it is imported into the clustered system.

2.5.7 Quorum disks

A quorum disk is a managed disk (MDisk) that contains a reserved area for use exclusively by the system. In the IBM Flex System V7000 Storage Node, any managed disks can be considered as a quorum candidate. The clustered system uses quorum disks to break a tie when exactly half the control canisters in the system remain active after a SAN failure. The IBM Flex System V7000 Storage Node dynamically assigns which quorum drive will be the active member (DQ).

The diagram in Figure 2-13 on page 61 shows a preferred quorum drive layout for a dual control enclosure clustered IBM Flex System V7000 Storage Node system.

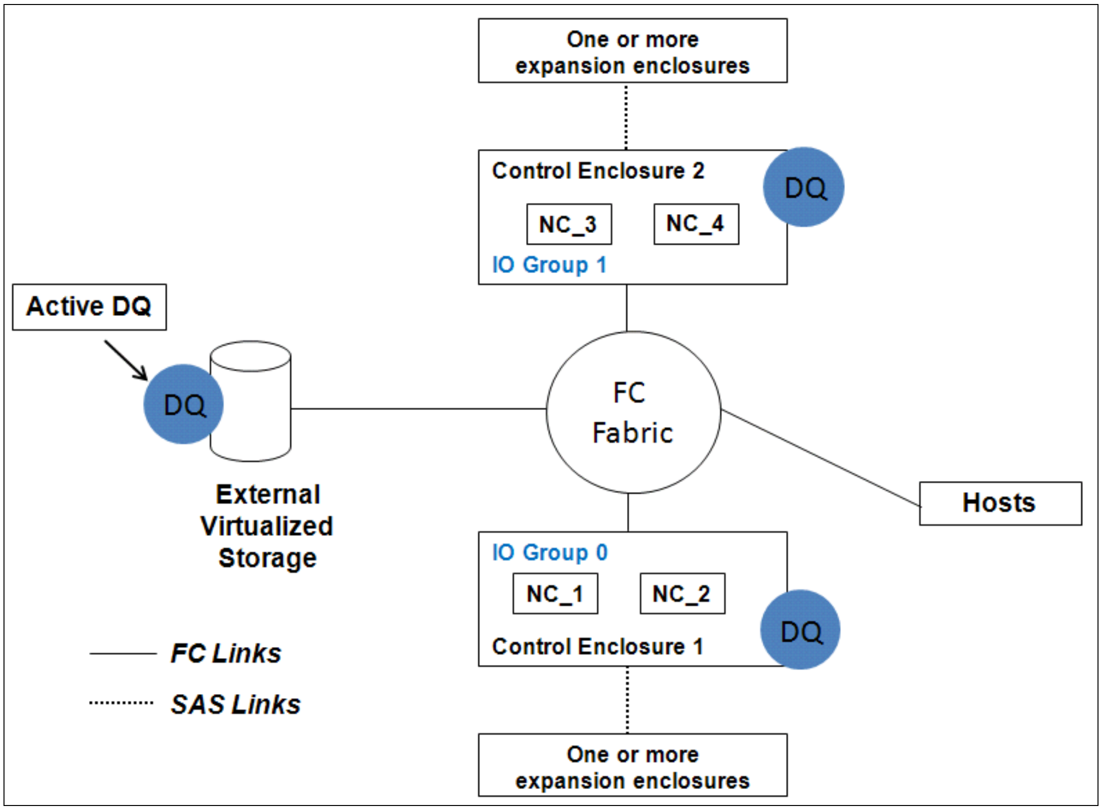


Figure 2-13 Preferred quorum drive layout

The clustered system automatically forms the quorum disk by taking a small amount of space from a managed disk (MDisk). It allocates space from up to three different MDisks for redundancy, although only one quorum disk is active.

If the environment has multiple storage systems, then to avoid the possibility of losing all of the quorum disks because of a failure of a single storage system, you should allocate the quorum disk on different storage systems. The preferred internal drives are the ones in the control enclosure and hot spares. However when an external virtualized storage system is being managed, it is preferred to have the active quorum disk located on it for better access when needed. It is possible to manage the quorum disks by using the CLI.

2.5.8 Storage pools

A storage pool is a collection of MDisks (up to 128) that are grouped together to provide capacity for the creation of the virtual volumes. All MDisks in the pool are split into extents with the same size. Volumes are then allocated out of the storage pool and are mapped to a host system.

> **Note:** All object names must begin with a alphabetic character and may not be numeric. The name can be a maximum of 63 characters. Valid characters are uppercase (A-Z), lowercase letters (a-z), digits (0 - 9), underscore (_), period (.), hyphen (-), and space; however, the names must not begin or end with a space.

MDisks can be added to a storage pool at any time to increase the capacity of the storage pool. MDisks can belong in only one storage pool and only MDisks in unmanaged mode can

be added to the storage pool. When an MDisk is added to the storage pool, the mode changes from unmanaged to managed and vice versa when you remove it.

Each MDisk in the storage pool is divided into a number of extents. The size of the extent is selected by the administrator at creation time of the storage pool and cannot be changed later. The size of the extent ranges from 16 MB up to 8 GB.

The extent size has a direct impact on the maximum volume size and storage capacity of the clustered system. A system can manage 4 million (4 x 1024 x 1024) extents. For example, a system with a 16 MB extent size can manage up to 16 MB x 4 MB = 64 TB of storage.

The effect of extent size on the maximum volume size is shown inTable 2-5, and lists the extent size and the corresponding maximum clustered system size.

Table 2-5 Maximum volume capacity by extent size

Extent size	Maximum volume capacity for normal volumes (GB)
16	2048 (2 TB)
32	4096 (4 TB)
64	8192 (8 TB)
128	16384 (16 TB)
256	32768 (32 TB)
512	65536 (64 TB)
1024	131072 (128 TB)
2048	262144 (256 TB)
4096	528288 (512 TB)
8192	1056576 (1,024 TB)

The effect of extent size on the maximum clustered system capacity is shown inTable 2-6.

Table 2-6 Extent size and effect on clustered system capacity

Extent size	Maximum storage capacity of cluster
16 MB	64 TB
32 MB	128 TB
64 MB	256 TB
128 MB	512 TB
256 MB	1 PB
512 MB	2 PB
1024 MB	4 PB
2048 MB	8 PB
4096 MB	16 PB
8192 MB	32 PB

Use the same extent size for all storage pools in a storage system, when they are supporting volume migration between two storage pools. If the storage pool extent sizes are not the same, use volume mirroring to copy volumes between storage pools, as described in Chapter 10, “Volume Mirroring and Migration” on page 423.

For most IBM Flex System V7000 Storage Nodes, a maximum of capacity of 1 PB is sufficient; and an extent size value of 256 MB should be used.

Default extent size: The GUI of IBM Flex System V7000 Storage Node sets a default extent size value of 256 MB when you define a new storage pool.

A storage pool can have a threshold warning set that automatically issues a warning alert when the used capacity of the storage pool exceeds the set limit.

Single-tiered storage pool

MDisks that are used in a single-tiered storage pool should have the following characteristics to prevent performance and other problems:

- They should have the same hardware characteristics, disk type, disk size and disk speeds (revolutions per minute (rpms)).
- They should have the same RAID type, RAID array size, and disk spindle count in the RAID grouping.
- The disk subsystems providing the MDisks must have similar characteristics, for example, maximum input/output operations per second (IOPS), response time, cache, and throughput.
- When possible, make all MDisks of the same size, and ensure that the MDisks provide the same number of extents per volume. If this configuration is not feasible, you need to check the distribution of the volumes’ extents in that storage pool.

Multitiered storage pool

A multitiered storage pool has a mix of MDisks with more than one type of disk tier attribute, for example, a storage pool containing a mix of generic_hdd AND generic_ssd MDisks.

A multitiered storage pool contains MDisks with different characteristics as opposed to the single-tiered storage pool. However, each tier should have MDisks of the same size and MDisks that provide the same number of extents.

A multitiered storage pool is used to enable automatic migration of extents between disk tiers using the IBM Flex System V7000 Storage Node Easy Tier function. Figure 2-14 on page 64 shows these components.

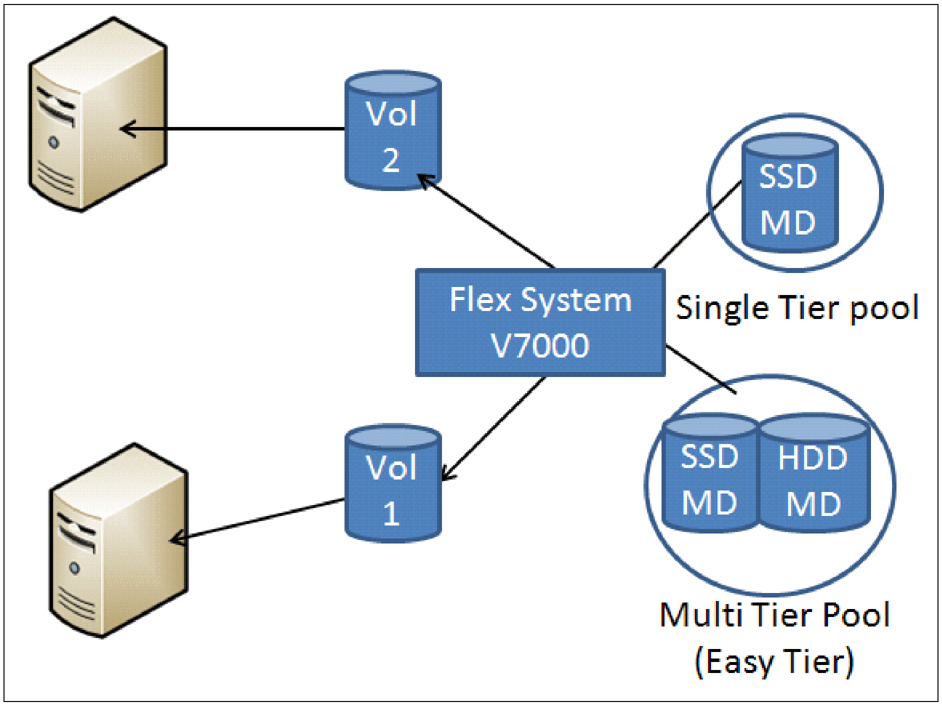


Figure 2-14 IBM Flex System V7000 Storage Node with single and multi-tier pools

2.5.9 Volumes

A volume is a virtual logical disk that is presented to a host system by the clustered system. In our virtualized environment, the host system has a volume mapped to it by IBM Flex System V7000 Storage Node. The IBM Flex System V7000 Storage Node translates this volume into a number of extents, which are allocated across MDisks. The advantage with storage virtualization is that the host is “decoupled” from the underlying storage, so the virtualization appliance can move the extents around without impacting the host system.

The host system cannot directly access the underlying MDisks in the same manner as it could access RAID arrays in a traditional storage environment.

There are three types of volumes:

- Striped Volume

  A striped volume is allocated using one extent from each MDisk at a time in the storage pool. This process continues until the space required for the volume has been satisfied. It is also possible to supply a list of MDisks to use.

- Sequential Volume

  A sequential volume is where the extents are allocated one after the other from one Mdisk. If there is not enough space on a single Mdisk the creation of the sequential volume fails. If a volume is required to be expanded the sequential volume is converted to a striped volume by policy when the expansion occurs

Figure 2-15 on page 65 shows examples of the striped and sequential volume types.

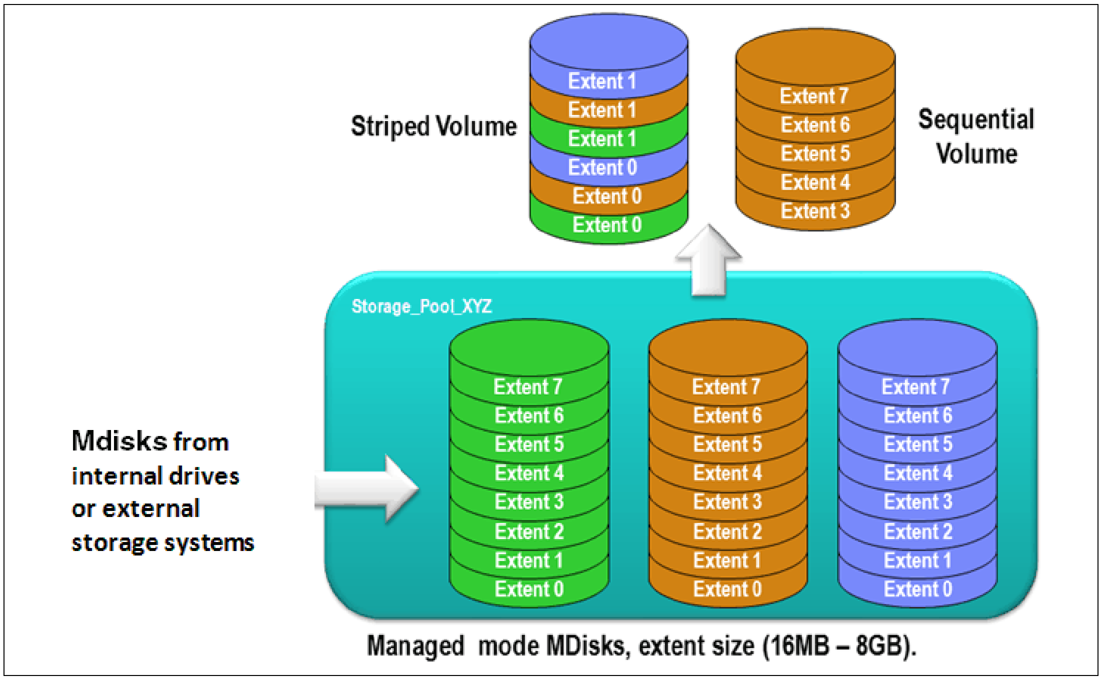


Figure 2-15 Striped and Sequential volumes

- Image mode

  Image mode volumes are special volumes that have a direct relationship with one MDisk. They are used to migrate existing data into and out of the IBM Flex System V7000 Storage Node.

  When the image mode volume is created, a direct mapping is made between extents that are on the MDisk and the extents that are on the volume. The logical block address (LBA) x on the MDisk is the same as the LBA x on the volume, which ensures that the data on the MDisk is preserved as it is brought into the clustered system (Figure 2-16 on page 66).

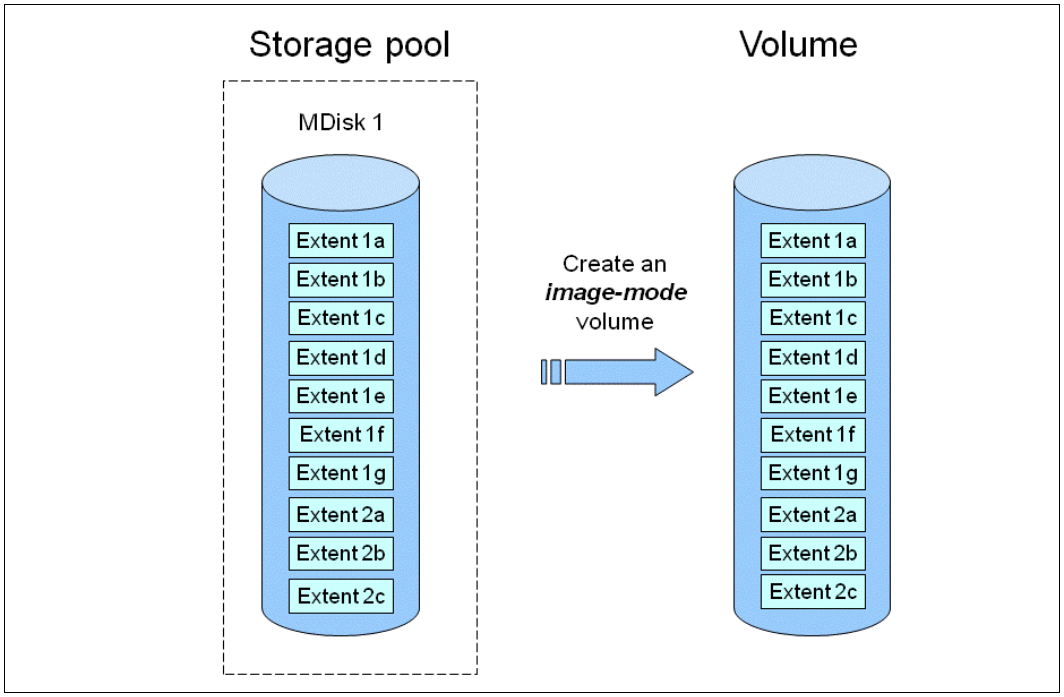


Figure 2-16 Image mode volume

Some virtualization functions are not available for image mode volumes, so it is often useful to migrate the volume into a new storage pool. After it is migrated, the MDisk becomes a managed MDisk.

If you add an MDisk containing data to a storage pool, any data on the MDisk is lost. Ensure that you create image mode volumes from MDisks that contain data before adding MDisks to the storage pools.

2.5.10 Thin-provisioned volumes

Volumes can be configured to either be *thin provisioned* or *fully allocated*. A thin-provisioned volume behaves with respect to application reads and writes as though they were fully allocated. When a volume is created, the user specifies two capacities: the real capacity of the volume and its virtual capacity.

The real capacity determines the quantity of MDisk extents that are allocated for the volume. The virtual capacity is the capacity of the volume reported to IBM Flex System V7000 Storage Node and to the host servers.

The real capacity is used to store both the user data and the metadata for the thin-provisioned volume. The real capacity can be specified as an absolute value or a percentage of the virtual capacity.

The thin provisioning feature can be used on its own to create overallocated volumes, or it can be used with FlashCopy. Thin-provisioned volumes can be used with the mirrored volume feature as well.

A thin-provisioned volume can be configured to *autoexpand*, which causes IBM Flex System V7000 Storage Node to automatically expand the real capacity of a thin-provisioned volume as its real capacity is used. Autoexpand attempts to maintain a fixed amount of unused real capacity on the volume. This amount is known as the *contingency capacity*.

The contingency capacity is initially set to the real capacity that is assigned when the volume is created. If the user modifies the real capacity, the contingency capacity is reset to be the difference between the used capacity and real capacity.

A volume that is created with a zero contingency capacity goes offline as soon as it needs to expand. A volume with a non-zero contingency capacity stays online until it has been used up.

Autoexpand does not cause the real capacity to grow much beyond the virtual capacity. The real capacity can be manually expanded to more than the maximum that is required by the current virtual capacity, and the contingency capacity is recalculated.

To support the autoexpansion of thin-provisioned volumes, the storage pools from which they are allocated have a configurable warning capacity. When the used free capacity of the group exceeds the warning capacity, a warning is logged. For example, if a warning of 80% has been specified, the warning is logged when 20% of the free capacity remains.

A thin-provisioned volume can be converted to a fully allocated volume using *volume mirroring* (and vice versa).

2.5.11 Mirrored volumes

IBM Flex System V7000 Storage Node provides a function called *volume mirroring*, which enables a volume to have two physical copies. Each volume copy can belong to a different storage pool and can be on different managed physical storage systems, which can help provide a better level of high-availability to a solution.

When a host system issues a write to a mirrored volume, IBM Flex System V7000 Storage Node writes the data to both copies. When a host system issues a read to a mirrored volume, IBM Flex System V7000 Storage Node requests it from the primary copy. If one of the mirrored volume copies is temporarily unavailable, the IBM Flex System V7000 Storage Node automatically uses the alternate copy without any outage for the host system. When the mirrored volume copy is repaired, IBM Flex System V7000 Storage Node resynchronizes the data.

A mirrored volume can be converted into a non-mirrored volume by deleting one copy or by splitting one copy to create a new non-mirrored volume.

The mirrored volume copy can be any type: image, striped, sequential, and thin provisioned or not. The two copies can be different volume types.

Using mirrored volumes can also assist with migrating volumes between storage pools that have different extent sizes and can provide a mechanism to migrate fully allocated volumes to thin-provisioned volumes without any host outages.

You can change the time-out value setting to either `latency`**,** which prioritizes low host latency (default) or `redundancy`, which prioritizes redundancy (longer timeout)

> **Unmirrored volumes:** An unmirrored volume can be migrated from one location to another by adding a second copy to the wanted destination, waiting for the two copies to synchronize, and then removing the original copy. This operation can be stopped at any time.

2.5.12 Easy Tier

Easy Tier is a performance function that automatically migrates or moves extents off a volume to, or from, SSD storage to HDD storage. Easy Tier monitors the host I/O activity and latency on the extent of all volumes with the Easy Tier function turned on in a multitiered storage pool over a 24-hour period. It then creates an extent migration plan based on this activity and then dynamically moves high activity or hot extents to a higher disk tier within the storage pool. It also moves extents' activity that has dropped off or cooled from the high tiered MDisk back to lower tiered MDisk. It should be understood that Easy Tier does not do any operations unless the results will have a level of positive value that makes the activity worth performing. In an environment with low total workload on a volume, even if the volume has a specific hot spot it may be judged to be too cool to justify the required operations to perform the move. Likewise, the downgrading of an extent will not take place until the space is needed for a hotter extent. Figure 2-17 shows the basic structure and concepts of this function.

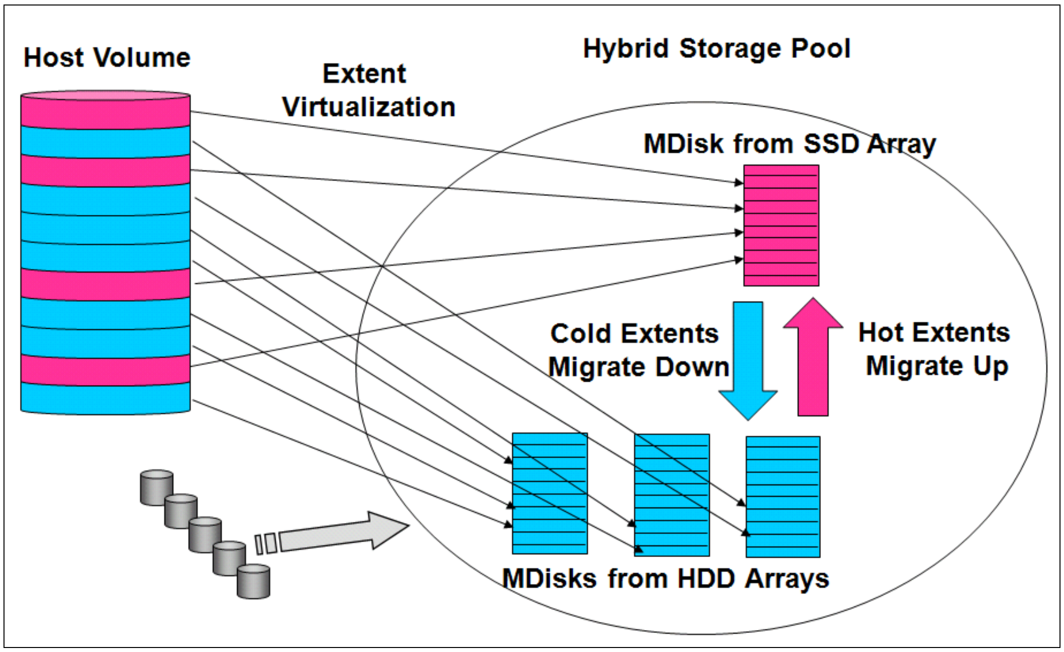


Figure 2-17 Easy Tier overview

The Easy Tier function may be turned on or off at the storage pool and volume level.

It is possible to demonstrate the potential benefit of Easy Tier in your environment before installing a solid-state disk. By turning on the Easy Tier function for a single level storage pool and the Easy Tier Function for the volumes within that pool, Easy Tier creates a migration report every 24 hours on the number of extents it would move if the pool was a multitiered pool. Easy Tier statistics measurement is enabled.

Using Easy Tier can make it more appropriate to use smaller storage pool extent sizes.

The usage statistics file can be offloaded from IBM Flex System V7000 Storage Node and then an IBM Storage Advisor Tool can be used to create a summary report from the data.

Contact your IBM representative or IBM Business Partner for more information about the Storage Advisor Tool. Refer to *Implementing the IBM Storwize V7000 V6.3*, SG24-7938 for more information about Easy Tier.

2.5.13 Real-time Compression (RTC)

With the IBM Flex System V7000 Storage Node there is a capability to create a volume as a compressed volume type. With this type of volume storage capacity needs can be lowered by as much as half. IBM Flex System V7000 Storage Node RTC function is based off of the same proven Random-Access Compression Engine (RACE) as the IBM RTC Appliances.

With compression, storage growth can be curbed and need for additional storage purchases can be delayed and spread out over greater periods of time. RTC dynamically works with active workloads now compresses the data while it is being processed the first time.

To implement RTC the volume must be created with compressed type selected. You cannot convert a volume from uncompressed to compressed after creation. However, a compressed volume can be a target of a volume mirror allowing the copying of the uncompressed volume to a compressed copy.

The RTC is a feature that is licensed on an enclosure basis for the IBM Flex System V7000 Storage Node.

RTC resource needs should be considered when planning for the use of compression. Resource requirements must come from someplace. Understanding the best balance of performance and compression is an important factor to consider when designing a mixed compressed environment.

RTC can be purchased to run on one control enclosure (IO Group) of a cluster and not for another allowing for shared environments to exist depending on need and providing environments that meet the needs.

To gain insight of the what should be expected a tool has been created which can be used to gather details and provide input on the workload type and pattern to show its level of compressibility. A tool called the Comprestimator can be used to analyze data gathered.

For more information on the usage and capabilities of this feature see *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859 at the following url:

http://w3.itso.ibm.com/redpieces/abstracts/redp4859.html?Open

2.5.14 iSCSI

iSCSI is an alternative means of attaching hosts to the IBM Flex System V7000 Storage Node. All communications with external back-end storage subsystems, or other IBM virtual storage systems must be done through a Fibre Channel or FCOE connection.

The iSCSI function is a *software function* that is provided by the IBM Flex System V7000 Storage Node code, not hardware.

In the simplest terms, iSCSI allows the transport of SCSI commands and data over a TCP/IP network, based on IP routers and Ethernet switches. iSCSI is a block-level protocol that encapsulates SCSI commands into TCP/IP packets and uses an existing IP network, instead of requiring expensive FC HBAs and a SAN fabric infrastructure.

A pure SCSI architecture is based on the client/server model. A client (for example, server or workstation) initiates read or write requests for data from a target server (for example, a data storage system).

Commands, which are sent by the client and processed by the server, are put into the Command Descriptor Block (CDB). The server runs a command, and completion is indicated by a special signal alert.

The major functions of iSCSI include encapsulation and the *reliable delivery* of CDB transactions between initiators and targets through the Internet Protocol network, especially over a potentially unreliable IP network.

The concepts of names and addresses have been carefully separated in iSCSI:

- An *iSCSI name* is a location-independent, permanent identifier for an iSCSI node. An iSCSI node has one iSCSI name, which stays constant for the life of the node. The terms *initiator name* and *target name* also refer to an iSCSI name.
- An *iSCSI address* specifies not only the iSCSI name of an iSCSI node, but also a location of that node. The address consists of a host name or IP address, a TCP port number (for the target), and the iSCSI name of the node. An iSCSI node can have any number of addresses, which can change at any time, particularly if they are assigned by way of Dynamic Host Configuration Protocol (DHCP). An IBM Flex System V7000 Storage Node control canister represents an iSCSI node and provides statically allocated IP addresses.

Each iSCSI node, that is, an initiator or target, has a unique iSCSI Qualified Name (IQN), which can have a size of up to 255 bytes. The IQN is formed according to the rules adopted for Internet nodes.

The iSCSI qualified name format is defined in RFC3720 and contains (in order) these elements:

- The string "`iqn`".
- A date code specifying the year and month in which the organization registered the domain or subdomain name used as the naming authority string.
- The organizational naming authority string, which consists of a valid, reversed domain or a subdomain name.
- Optionally, a colon (:), followed by a string of the assigning organization's choosing, which must make each assigned iSCSI name unique.

For IBM Flex System V7000 Storage Node the IQN for its iSCSI target is specified as:

```
iqn.1986-03.com.ibm:2145.<clustername>.<nodename>
```

On a Windows server, the IQN, that is, the name for the iSCSI initiator, can be defined as:

```
iqn.1991-05.com.microsoft:<computer name>
```

The IQNs can be abbreviated by using a descriptive name, known as an *alias*. An alias can be assigned to an initiator or a target. The alias is independent of the name and does not need to be unique. Because it is not unique, the alias must be used in a purely informational way. It cannot be used to specify a target at login or used during authentication. Both targets and initiators can have aliases.

An iSCSI name provides the correct identification of an iSCSI device irrespective of its physical location. Remember, the IQN is an *identifier,* not an *address*.

Changing names: Before changing system or node names for an IBM Flex System V7000 Storage Node clustered system that has servers connected to it using SCSI, be aware that because the system and node name are part of the IQN for the IBM Flex System V7000 Storage Node you can lose access to your data by changing these names. The IBM Flex System V7000 Storage Node GUI shows a specific warning, but the CLI does *not*.

The iSCSI session, which consists of a *login phase* and a *full feature phase*, is completed with a special command.

The login phase of the iSCSI is identical to the FC port login process (PLOGI). It is used to adjust various parameters between two network entities and to confirm the access rights of an initiator.

If the iSCSI login phase is completed successfully, the target confirms the login for the initiator; otherwise, the login is not confirmed and the TCP connection breaks.

As soon as the login is confirmed, the iSCSI session enters the full feature phase. If more than one TCP connection was established, iSCSI requires that each command / response pair goes through one TCP connection. Thus, each separate read or write command is carried out without the necessity to trace each request for passing separate flows. However, separate transactions can be delivered through separate TCP connections within one session.

For further details about configuring iSCSI, see Chapter 11, “SAN Connections and Configuration” on page 431.

2.5.15 Fiber Channel over Ethernet (FCoE)

Fiber Channel over Ethernet (FCoE) is a standard specified by ANSI T11 committee within the FC-BB-5 for enabling the transmission of FC protocol and data across an ethernet network. As shown in Figure 2-18, the new environment supports both FCoE and TCP/IP traffic to be able to share a common ethernet network.

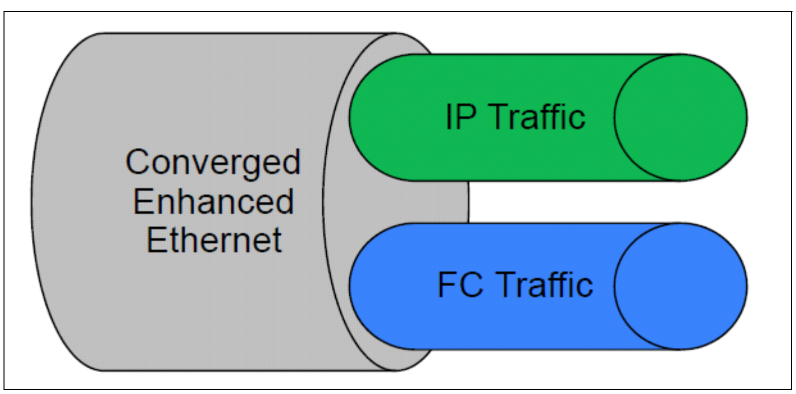


Figure 2-18 New enhanced ethernet environment support

Table 2-7 on page 72 shows the differences between the use of FCoE and iSCSI for transfer environments.

Table 2-7 FCoE and iSCSI differences

FCoE	iSCSI
Local-area, lossless links, no routing allowed	Allows many hops, lossy connections and high latency
Simple encapsulation of Fibre Channel	Substantial complexity on top of TCP
Low overhead – similar to Fibre Channel	Overhead varies - typically higher
Storage administrators know Fiber Channel well	TCP/iP well understood by most IT staff
Frame loss can quickly become catastrophic	Frame loss recovery built into protocol stack

FCoE allows for fewer network adapters to be required as both protocols can share the same adapter, reducing hardware cost and freeing up bus slots in hosts.

For in depth details and recommendations for using these protocols see *Storage and Network Convergence Using FCoE and iSCSI*, SG24-7986 available at:

http://w3.itso.ibm.com/abstracts/sg247986.html?Open

2.6 Advanced copy services

IBM Flex System V7000 Storage Node supports the following Copy services:

- FlashCopy
- Synchronous Remote Copy
- Asynchronous Remote Copy

2.6.1 FlashCopy

FlashCopy makes a copy of a source volume on a target volume. The original content of the target volume is lost. After the copy operation has started, the target volume has the contents of the source volume as they existed at a single point in time. Although the copy operation takes time, the resulting data at the target appears as though the copy was made instantaneously.

FlashCopy is sometimes described as an instance of a time-zero (T0) copy or a point in time (PiT) copy technology.

FlashCopy can be performed on multiple source and target volumes. FlashCopy permits the management operations to be coordinated so that a common single point in time is chosen for copying target volumes from their respective source volumes.

IBM Flex System V7000 Storage Node also permits multiple target volumes to be FlashCopied from the same source volume. This capability can be used to create images from separate points in time for the source volume, and to create multiple images from a source volume at a common point in time. Source and target volumes can be thin-provisioned volumes.

Reverse FlashCopy enables target volumes to become restore points for the source volume without breaking the FlashCopy relationship and without waiting for the original copy operation to complete. IBM Flex System V7000 Storage Node supports multiple targets and thus multiple rollback points.

Most clients aim to integrate the FlashCopy feature for point in time copies and quick recovery of their applications and databases. IBM Support is provided by Tivoli Storage FlashCopy Manager, which is described at the following website:

`http://www-01.ibm.com/software/tivoli/products/storage-flashcopy-mgr/`

You can read a detailed description about the FlashCopy copy services in 9.2, “FlashCopy” on page 342.

2.6.2 IBM Flex System V7000 Remote Mirroring Software

The IBM Flex System V7000 Remote Mirroring provides both Metro and Global mirroring capabilities between the IBM Flex System V7000 Storage Nodes, or between IBM Flex System V7000 Storage Nodes and IBM Storwize V7000s or IBM SAN Volume Controllers (SVC). This capability means customers have greater flexibility in their expanding environments using Flex System V7000, Storwize V7000, and SVC with the ability now to remote mirror from one system to the other. With the wide variety of storage systems that can be managed under any of these systems, the options available for replicating data between a number of storage systems is multiplied. Remote deployments for disaster recovery for current SVC and Storwize V7000 environments can easily be fitted with the Flex System V7000 or vice versa.

The Copy Services layer sits above and operates independently of the function or characteristics of the underlying disk subsystems used to provide the storage resources to the IBM Flex System V7000 Storage Node. Figure 2-19 shows an example of possible copy service relationships (they must be at Version 6.4 or later).

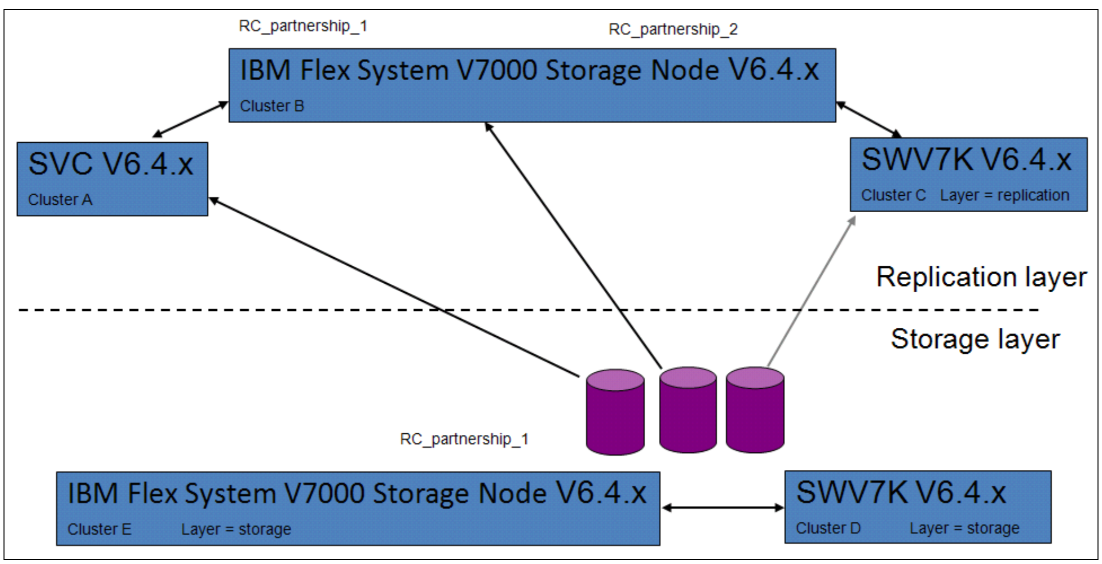


Figure 2-19 Example of possible copy services relationships

With Metro Mirroring the IBM Flex System V7000 Remote Mirroring provides synchronous data replication at distances less than 300Km. This capability is supported across either an FC or FCoE SAN fabric infrastructure. Mirroring over iSCSI connections is not supported at this time.

Global Mirroring provides for long distances of greater than 300 Km and is only supported over an FC SAN infrastructure. Tunneling over a WAN network is frequently used for the greater distance needs.

Customers who wish to use the global mirror capability with Flex System V7000 on a low bandwidth link between sites can do so with the use of the low bandwidth remote mirroring. This capability provides options to help administrators balance network bandwidth requirements and RPO times for applications, helping reduce operation costs for disaster recovery solutions. Remote mirroring supports higher RPO times by allowing the data at the disaster recovery site to get further out of sync with the production site if the communication link limits replication, and then approaches synchronicity again when the link is not as busy. This low bandwidth remote mirroring uses space-efficient FlashCopy targets as sources in remote copy relationships to increase the time allowed to complete a remote copy data cycle.

Refer to 9.3, “Remote Copy” on page 380 for more information about the Remote Copy services.

2.6.3 Remote Copy licensing

To authorize use of remote copy capabilities of the IBM Flex System V7000, you must purchase a feature license for IBM Flex System V7000 Remote Mirroring Software. This license is based on the quantity of enclosures that are authorized for the feature.

To determine the number of authorized enclosures needed you should count all enclosures on the system that are managed by the IBM Flex System V7000 Disk System, including each internal enclosure licensed with the IBM Flex System V7000 Base Software and each external enclosure licensed with the IBM Flex System V7000 External Virtualization Software. This should be done at both the ends of the mirroring relationship. The location with the least number of members is the location that defines the number of licenses needed for each location site.

For example, if you have site A with a DS5020 consisting of three enclosures managed by an IBM Flex System V7000 consisting of one control enclosure and one expansion enclosure mirroring to a partner at site B with an IBM Flex System V7000 consisting of one control enclosure and two expansion enclosures and no external virtual storage system attached, then in order to authorize remote mirroring for this configuration you will need a license to the IBM Flex System V7000 Remote Mirroring Software, for a quantity of six enclosures; three for site A and three for site B. If however, in the future you were to grow your site B system to include more enclosures or an external virtualized system you would then need to re-evaluate the licensing to add the appropriate number for the new configuration of the site with the fewest enclosures managed.

Remote mirroring Flex System V7000 with an SVC in the mix, because of a difference in metrics, SVC mirroring can be licensed for a subset of the total storage virtualized. However, the IBM Flex System V7000 and the IBM Storwize V7000 still must used the above model for their licensing.

2.6.4 Synchronous / Asynchronous Remote Copy

The general application of Remote Copy seeks to maintain two copies of data. Often, the two copies are separated by distance, but not always.

The Remote Copy can be maintained in one of two modes: synchronous or asynchronous.

With the IBM Flex System V7000 Storage Node, Metro Mirror and Global Mirror are the IBM branded terms for the functions that are synchronous Remote Copy and asynchronous Remote Copy.

Synchronous Remote Copy ensures that updates are committed at both the primary and the secondary before the application considers the updates complete; therefore, the secondary is fully up to date if it is needed in a failover. However, the application is fully exposed to the latency and bandwidth limitations of the communication link to the secondary. In a truly remote situation, this extra latency can have a significant adverse effect on application performance.

Special configuration guidelines exist for SAN fabrics that are used for data replication. It is necessary to consider the distance and available bandwidth total network links to determine the appropriate method to use. See 9.3, "Remote Copy" on page 380 for details on planning, configuring and using Remote Copy for replication functions.

With the Global Mirror method there is a design option that assists with low bandwidth for theIBM Flex System V7000 Storage Node and the other IBM virtual storage systems that aremirroring to it running 6.4.1 code level. This option uses change volumes associated with the primary and secondary volumes. These volumes are used to record changes to the remote copy volume, the FlashCopy relationship that exists between the secondary volume and the change volume, and between the primary volume and the change volume. This function is called Global Mirror cycling mode. Figure 2-20 shows an example of this function where you can see the relationship between volumes and change volumes.

> **Note:** Though this feature was added in the 6.3 code level only 6.4.1code level is supported to mirror to the IBM Flex System V7000 Storage Node.

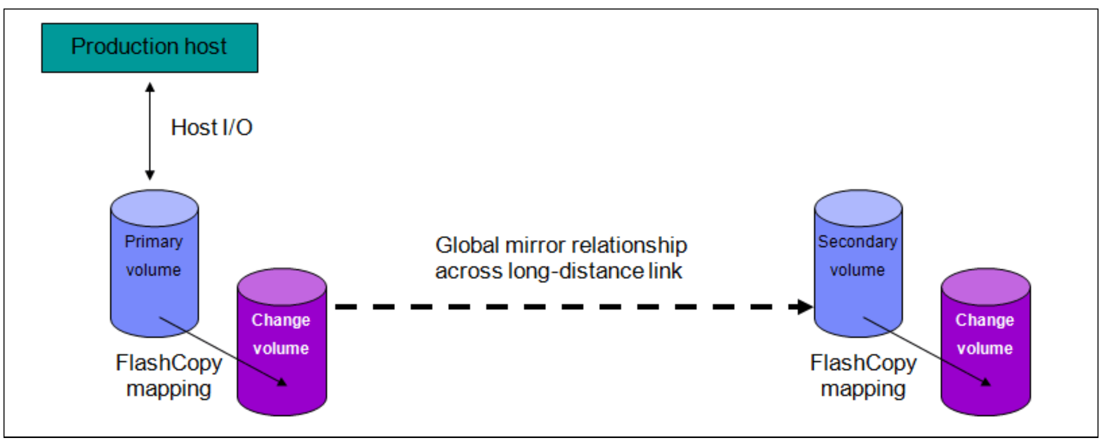


Figure 2-20 Global Mirror with change volumes

In asynchronous Remote Copy, the application acknowledges that the write is complete before the write is committed at the secondary. Hence, on a failover, certain updates (data) might be missing at the secondary. The application must have an external mechanism for recovering the missing updates, if possible. This mechanism can involve user intervention. Recovery on the secondary site involves starting the application on this recent "backup" and then rolling forward or backward to the most recent commit point.

2.6.5 Copy Services configuration limits

In Table 2-8 on page 76, we describe the Copy Services configuration limits. For the most up-to-date list of these limits, see the following website:

http://www.ibm.com/support/docview.wss?uid=ssg1S1003702&myns=s028&mynp=familyind5402112&mync=E

Table 2-8 Copy Services configuration limits

Properties	Maximum number	Note
Remote Copy (Metro Mirror and Global Mirror) relationships per clustered system	4096	This configuration can be any mix of Metro Mirror and Global Mirror relationships.
Remote Copy relationships per consistency group	4096	No additional limit is imposed beyond the Remote Copy relationships per clustered system limit
Remote Copy consistency groups per clustered system	256	-
Total Metro Mirror and Global Mirror volume capacity per I/O group	1024 TB	This limit is the total capacity for all master and auxiliary volumes in the I/O group.
FlashCopy mappings per clustered system	4096	-
FlashCopy targets per source	256	-
Cascaded Incremental FlashCopy maps	4	A volume can be the source of up to four incremental FlashCopy maps. If this number of maps is exceeded, then the FlashCopy behavior for that cascade becomes non-incremental.
FlashCopy mappings per consistency group	512	-
FlashCopy consistency groups per clustered system	127	-
Total FlashCopy volume capacity per I/O group	1024 TB	4096 for a full four node clustered system with four I/O groups

2.7 Management and support tools

The IBM Flex System V7000 Storage Node system can be managed through the IBM Flex System Management Node (FSM), or by using the native management software that runs in the hardware itself.

The FSM simplifies storage management in the following ways:

- Centralizes the management of storage network resources with IBM storage management software.
- Provides greater synergy between storage management software and IBM storage devices.
- Reduces the number of servers that are required to manage your software infrastructure.
- Provides higher-level functions.

2.7.1 IBM Assist On-site and remote service

The IBM Assist On-site tool is a remote desktop-sharing solution that is offered through the IBM website. With it, the IBM service representative can remotely view your system to troubleshoot a problem.

You can maintain a chat session with the IBM service representative so that you can monitor this activity and either understand how to fix the problem yourself or allow the representative to fix it for you.

To use the IBM Assist On-site tool, the SSPC or master console must be able to access the Internet. The following website provides further information about this tool:

http://www.ibm.com/support/assistonsite/

When you access the website, you sign in and enter a code that the IBM service representative provides to you. This code is unique to each IBM Assist On-site session. A plug-in is downloaded on to your SSPC or master console to connect you and your IBM service representative to the remote service session. The IBM Assist On-site contains several layers of security to protect your applications and your computers.

You can also use security features to restrict access by the IBM service representative.

Your IBM service representative can provide you with more detailed instructions for using the tool.

2.7.2 Event notifications

IBM Flex System V7000 Storage Node can use Simple Network Management Protocol (SNMP) traps, syslog messages, and a Call Home email to notify you and the IBM Support Center when significant events are detected. Any combination of these notification methods can be used simultaneously.

Each event that IBM Flex System V7000 Storage Node detects is assigned a notification type of Error, Warning, or Information. You can configure IBM Flex System V7000 Storage Node to send each type of notification to specific recipients.

2.7.3 SNMP traps

SNMP is a standard protocol for managing networks and exchanging messages. IBM Flex System V7000 Storage Node can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that IBM Flex System V7000 Storage Node sends. You can use the management GUI or the IBM Flex System V7000 Storage Node command-line interface to configure and modify your SNMP settings.

You can use the Management Information Base (MIB) file for SNMP to configure a network management program to receive SNMP messages that are sent by the IBM Flex System V7000 Storage Node. This file can be used with SNMP messages from all versions of IBM Flex System V7000 Storage Node Software.

2.7.4 Syslog messages

The syslog protocol is a standard protocol for forwarding log messages from a sender to a receiver on an IP network. The IP network can be either IPv4 or IPv6. IBM Flex System V7000 Storage Node can send syslog messages that notify personnel about an event. IBM

Flex System V7000 Storage Node can transmit syslog messages in either expanded or concise format. You can use a syslog manager to view the syslog messages that IBM Flex System V7000 Storage Node sends. IBM Flex System V7000 Storage Node uses the User Datagram Protocol (UDP) to transmit the syslog message. You can use the management GUI or the IBM Flex System V7000 Storage Node command-line interface to configure and modify your syslog settings.

2.7.5 Call Home email

The Call Home feature transmits operational and error-related data to you and IBM through a Simple Mail Transfer Protocol (SMTP) server connection in the form of an event notification email. When configured, this function alerts IBM service personnel about hardware failures and potentially serious configuration or environmental issues. You can use the call home function if you have a maintenance contract with IBM or if the IBM Flex System V7000 Storage Node is within the warranty period.

To send email, you must configure at least one SMTP server. You can specify as many as five additional SMTP servers for backup purposes. The SMTP server must accept the relaying of email from the IBM Flex System V7000 Storage Node clustered system IP address. You can then use the management GUI or the command-line interface to configure the email settings, including contact information and email recipients. Set the reply address to a valid email address. Send a test email to check that all connections and infrastructure are set up correctly. You can disable the Call Home function at any time using the management GUI or the command-line interface as well.

2.8 Useful references from Storwize V7000 websites

Due to the common design of the software and procedures of the IBM Storwize V7000 and the IBM Flex System V7000 Storage Node the following information is provided for additional references.

- IBM Flex System Information Center:

  http://publib.boulder.ibm.com/infocenter/flexsys/information/index.jsp

- IBM Flex System V7000 Storage Node Information Center:

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_product_page.html

- IBM Flex System V7000 Storage Node - Learning Modules
- http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.learning/site_learning_tutorials.html
- IBM Flex System V7000 Storage Node Installation Guide (GC27-4203)

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_bkmap_quickinstbk.pdf

- IBM Storwize V7000 Expansion Enclosure Installation Guide, Machine type 2076 (GC27-4234)

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/tbrd_bkmap_install_2076.pdf

- IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide (GC27-4205)

http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_bkmap_trblshoot.pdf

- IBM Flex System V7000 Storage Node Support home:

  http://www.ibm.com/support/entry/portal/overview/hardware/puresystems/pureflex_system/storage_node/flex_system_v7000

- The IBM Storwize V7000 Storage Support page is at the following website

  http://www.ibm.com/storage/support/storwize/v7000

- The IBM Storwize V7000 Storage home page is at the following website

  http://www-03.ibm.com/systems/storage/news/center/storwize_v7000/index.html

- The Storwize V7000 Storage Supported hardware list is at the following website:

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003908

- The Storwize V7000 Storage Configuration Limit and Restrictions are at the following website:

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003902

- The Storwize V7000 Storage Online documentation is at the following website:

  http://www-947.ibm.com/support/entry/portal/Documentation/Hardware/System_Storage/Disk_systems/Mid-range_disk_systems/IBM_Storwize_V7000_%282076%29

- The Storwize V7000 Storage Online Information Center is at the following website:

  http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp

- You can see IBM Redbooks publications about both the IBM Flex System and IBM Storwize V7000 Storage at the following website:

  http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=v7000

2.8.1 IBM virtual storage learning videos on YouTube

Some interesting videos are available on YouTube that describe the IBM Storwize V7000 Storage. With the common software design shared with this system and the IBM Flex System V7000 Storage Node many of these can be used to help develop a familiarity with the terminology and many of the GUI and management processes. These are available at the URLs shown in Table 2-9.

Table 2-9 Videos available on YouTube

Description of the video	Link WEB
IBM Storwize V7000 Storage Virtualization Terminology Overview	http://www.youtube.com/watch?v=I2rzt3m2gP0
IBM Storwize V7000 Interface tour	http://www.youtube.com/watch?v=FPbNRs9HacQ
IBM Storwize V7000 Volume management	http://www.youtube.com/watch?v=YXeKqH8Sd9o
IBM Storwize V7000 Migration	http://www.youtube.com/watch?v=dXxnUN6dk74
IBM Storwize V7000 Introduction to FlashCopy	http://www.youtube.com/watch?v=MXWgGWjBzG4
IBM Storwize V7000 Introduction to Remote Copy 1	http://www.youtube.com/watch?v=koFYm--gnEc
IBM Storwize V7000 Introduction to Remote Copy 2	http://www.youtube.com/watch?v=kMOFFS3NYas

Description of the video	Link WEB
IBM Storwize V7000 Rapid Application Storage Solution - Tivoli Storage FlashCopy Manager	http://www.youtube.com/watch?v=1zgEljRVDSc
VMware data protection with Storwize V7000	http://www.youtube.com/watch?v=vec0ap-qwbA
IBM SAN Volume Controller and Storwize V7000 Performance Panel Sped-up! (HD)	http://www.youtube.com/watch?v=7noC71tLkWs

3

Systems Management

This chapter provides an overview of storage management for IBM Flex System V7000 Storage Node from the IBM Flex System Chassis Management Module and the IBM Flex System Manager Node, and shows how to use the navigation tools.

While there are several ways to manage the various components of the IBM Flex Systems, a discussion about all these options is beyond the scope of this publication. In this chapter, we will cover the management options related to IBM Flex System V7000 Storage Node and the I/O modules. For detailed information regarding other management options, refer to *Implementing Systems Management of IBM PureFlex System*, SG24-8060.

Here is a list of topics discussed in this chapter:

3.1 System Management Overview

The IBM PureFlex System is designed to improve efficiency in system management. The advanced automation of configuration and management, enhanced and user-friendly interface, cross-resource integration, easier monitoring and problem management, policy-based expert patterns, and all other functions and features are designed towards providing a simplified yet dynamic management experience to the user.

System Management can be performed by using easy-to-use integrated platform management tools (e.g. CMM and IMMv2.) as well as other popular system management software like IBM Tivoli and IBM Director, and also third-party enterprise management tools (e.g. CA, BMC and HP). Due to the flexibility of an IBM Flex System, many management features and functions like security, monitoring, provisioning and automation, can be plugged easily into an existing IT infrastructure.

Within an IBM Flex System Enterprise Chassis, each node has a management controller that communicates via the CMM-enabled 1GbE private management network allowing centralized management capability. Different chassis components supported in the IBM Flex System Enterprise Chassis implement different management controllers at node-level as well.

The following set of hardware management devices and interfaces are commonly used for managing an IBM Flex System:

- IBM Flex System Manager (FSM)
- IBM Flex System Chassis Management Module (CMM)
- Switch modules (IO modules)
- Compute node management processors
  - Integrated Management Module (IMMv2) for intel-based servers
  - Flexible Service Processor (FSP) for POWER-based servers

IBM Flex System Manager (FSM)

FSM is a highly integrated management device that offers single-system management across physical and virtual resources. It provides complete control of IBM Flex System components and features. It offers:

- Virtualization management
- Resource allocation and control
- Network and storage control

More information regarding IBM Flex System Manager can be found at 3.3.1, “Overview of FSM” on page 110

IBM Flex System Chassis Management Module (CMM)

CMM is the default integrated platform for managing a single IBM Flex System Enterprise Chassis. It occupies one of the two chassis management bays on rear of chassis. A second module can be installed for redundancy.

More information regarding IBM Flex System Chassis Management Module can be found at 3.2.1, “Overview of IBM Flex System Chassis Management Module” on page 88.

I/O Modules (Switch modules)

I/O modules contain the internal management interfaces and Inter-Integrated Circuit (I2C) for management. They can be managed by the IBM Flex System Manager, IBM Flex System

Chassis Management Module, command line interface (CLI), web interface, IBM System Network Element Manager (SNEM), or any third-party SNMP-based management tool.

Compute Node Management Processors

There are two types of compute nodes in an IBM Flex System:

- For IBM x86 (intel-based) compute nodes, the management interface is IMMv2, which is a powerful management interface providing control at node-level.
- For POWER compute node, Flexible Service Processor(FSP) provides a serial-over-lan (SOL) interface for performing common management tasks on the node.

Table 3-1 gives an overview of system management capabilities of these hardware management devices.

Table 3-1 System management

IBM Flex System Manager	IBM Flex System Chassis Management Module	IMMv2 / Flexible Service Processor	Network Switch
► Remote access via browser ► Multi-Chassis Rack Management ► Configuration and Updates ► Network Configuration Manger ► Service & Support Manager ► Virtualization Management (VM Control) ► Active Energy Manager ► Local Console for Chassis Management ► Remote control for Compute node(s)	► Aggregation point for chassis ► Integrated Remote Chassis Management ► Power and Thermal Management ► Internal Chassis Network with Management packet forwarding ► Web, CLI, SNMP, CIM-XML ► CMM redundancy (optional additional CMM) ► RAS, FFDC, Service &Support, Chassis Event Logs	► Remote Presence (CLI/video/USB) ► Power and Thermal Management. ► Inventory, Fault detection ► Update Management ► Configuration Management ► RAS (reliability, availability, serviceability), FFDC (First-failure data capture), Logs	► Power, Thermal sensing ► Update Management ► Inventory, Fault detection ► Configuration Management ► RAS, FFDC (First-failure data capture), Logs ► SNMP, CLI, Web, other proprietary protocols

3.1.1 IBM Flex System Storage Management

The IBM Flex System offers strong storage management capabilities that allow you to gain advanced functionality with storage nodes in your system while also taking advantage of your existing storage infrastructure through advanced virtualization.

The advanced Storage Management features in an IBM Flex System include:

- **Automated discovery and inventory** - Physical seating of Storage Node triggers discovery and inventory, automatically adding it to IBM Flex System Manager
- **Intelligent storage provisioning** - Integrated storage allocation
- **Automated firmware update** - Update Manager provides firmware level compliance monitoring and management
- **Integrated event monitoring** - Seamless visibility to events and status
- **Automatic credential exchange** - Storage Node's initial credentials are set via IBM Flex System Manager

IBM Flex System V7000 Storage Node

IBM Flex System V7000 Storage Node, is physically and logically integrated into an IBM Flex System. IBM Flex System V7000 Storage Node is designed on the industry-leading storage

virtualization and efficiency capabilities of IBM Storwize V7000 by simplifying and speeding deployment, and providing greater integration of server and storage management to automate and streamline provisioning for greater responsiveness to business needs and lower overall cost.

IBM Flex System V7000 Storage Node is an Integrated storage system with advanced storage efficiency capabilities, which include Thin provisioning, FlashCopy, Easy Tier, Real-time Compression, Non disruptive Migration. It offers external virtualization for rapid data center integration, Metro and Global Mirror for multi-site recovery, scalability up to 240 small form-factor (SFF) drives (HDD and SSD), and Clustered systems support up to 960 drives.

IBM Flex System V7000 Storage Node simplifies storage administration with a single user interface for all your storage with a management console that is integrated with the comprehensive management system. The enhanced management capabilities allow you to virtualize third-party storage with non-disruptive migration of the existing storage infrastructure. Intelligent tiering provides the advantage of balancing high performance and cost for your storage needs. It also protects data for business continuity and disaster recovery with built-in local and remote replication and snapshot functions.

IBM Flex System V7000 Storage Node enhancements include physical chassis plug-and-play integration, automated deployment and discovery, integration into Flex System Manager Chassis map, and FCoE optimized offering in addition to FC and iSCSI.

For storage management, the IBM Flex System V7000 Storage Node system can have two management IP addresses.

- One for the 1 Gb/s internal management network (required)
- One on the first 10 Gb/s Ethernet port in the configuration node canister (optional)

3.1.2 Storage Management Interfaces

The IBM Flex System storage infrastructure contains a variety of management interfaces that can be classified into four distinct levels based on their level of depth and detail in management capabilities:

Starting Level - IBM Flex System Manager Chassis Map

The IBM Flex System Manager offers a Chassis Map, which is the single point of management and starting point for IBM Flex System. It allows for physical navigation through the entire system, as shown in Figure 3-1.

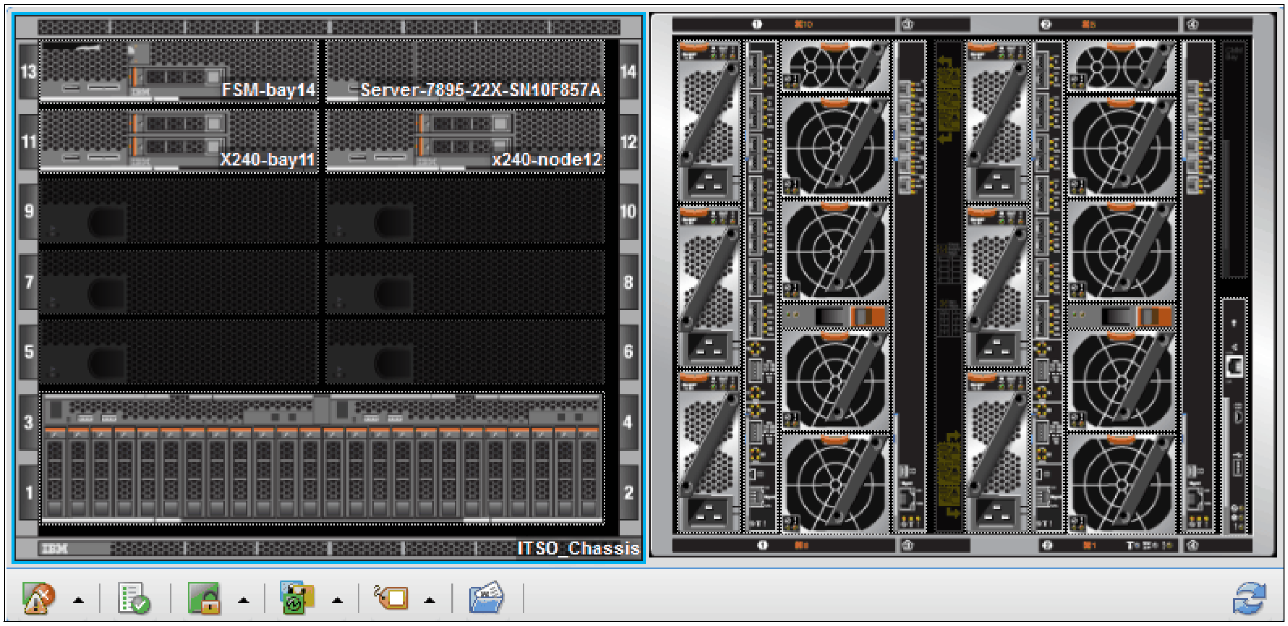


Figure 3-1 FSM chassis Map

Upper Level - IBM Flex System Manager Storage Control

Higher level datacenter management can be done using IBM Flex System Manager Storage Control feature. It provides integrated management with a Systems and Storage perspective. Using Storage Control, you can manage storage across your data center connected to IBM Flex System, and also map connections between servers, switches and storage. Storage Control is shown in Figure 3-2 on page 86.

Figure 3-2 IBM IBM Flex System Manager Storage control

Detailed Level - IBM Flex System V7000 Storage Node User Interface

This interface is inherent in every IBM Flex System V7000 Storage Node. It allows for detailed disk setup (RAID arrays, LUNs). All the advanced disk features like Easy Tier, Mirroring, Clustering, and Virtualization can be configured using the IBM Flex System V7000 Storage Node User Interface. Figure 3-3 on page 87 shows the IBM Flex System V7000 Storage Node interface.

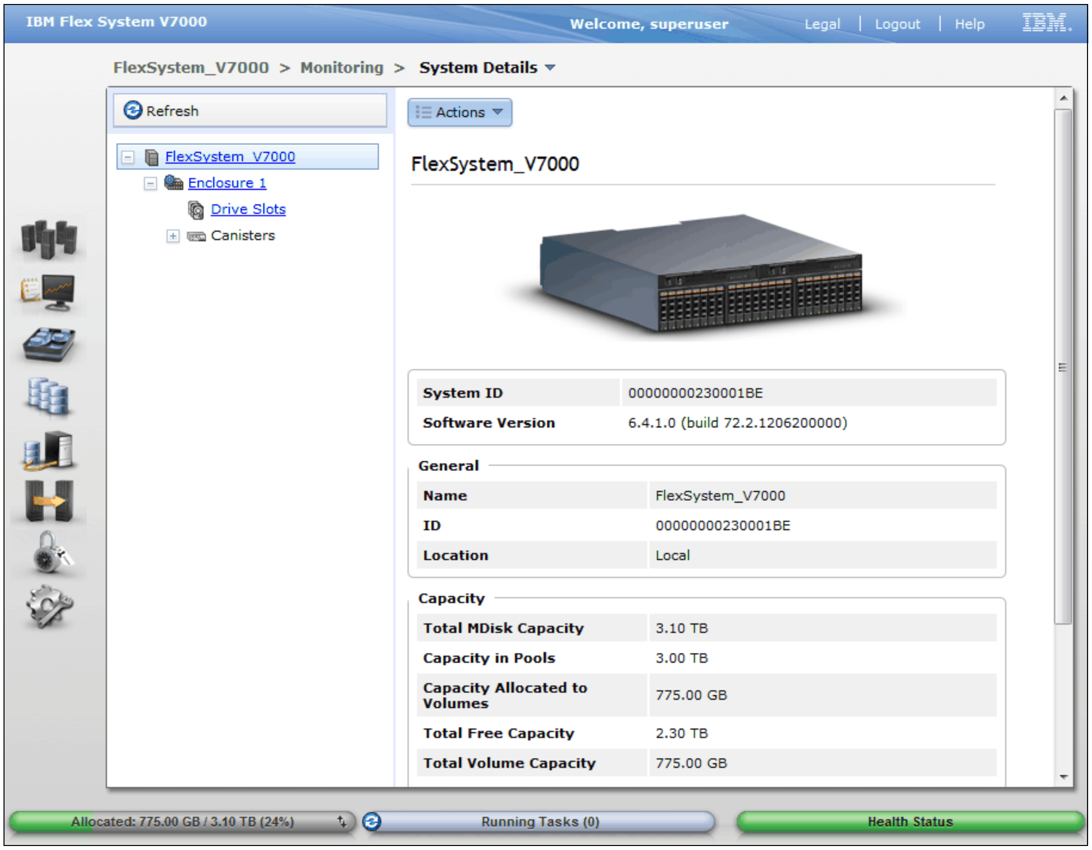


Figure 3-3 IBM Flex System V7000 GUI - System Details

For further details on management of an IBM Flex System V7000 Storage Node, refer to Chapter 5, “IBM Flex System V7000 Storage Node Graphical user interface overview” on page 175.

3.2 IBM Flex System Chassis Management Module (CMM)

The IBM Flex System Chassis Management Module (CMM) is a hot-swap module that configures and manages all installed chassis components. The chassis comes with one IBM Flex System Chassis Management Module in the CMM bays.

The IBM Flex System Chassis Management Module auto-detects and manages the hardware of the entire IBM Flex System Enterprise Chassis. It communicates with the management controller in each node (IMMv2 in intel processor-based computer nodes, FSP in power Power7 processor-based nodes), along with the management applications of storage node and I/O modules to provide centralized system monitoring and management.

3.2.1 Overview of IBM Flex System Chassis Management Module

A IBM Flex System Chassis Management Module (CMM) gives control over the IBM Flex System at the chassis level, simplifying installation and management of your installation.

The CMM is central to the management of the chassis, and is required in the IBM Flex System Enterprise Chassis. One CMM comes standard with the chassis. There are two CMM bays in the chassis, giving the user an option to add a second CMM for redundancy.

Through an embedded firmware stack, the CMM implements functions to monitor, control, and provide external user interfaces to manage all chassis resources. Key functions of the IBM Flex System Chassis Management Module are:

- Single-chassis management
- Initial setup configuration wizard from http/s interface
- Detailed system health status for components, along with suggested recovery actions
- Chassis level Hardware inventory management
- Chassis Configuration and Validation
- System and component Status information
- Component LEDs states and controls.
- Event handling and Notification
- Hardware access states
- Fan status and control
- Thermal management and monitoring
- Power management and control
- Switch modules - monitor and Control access
- Security features and policies
- Service data collection and call home services
- Configuration backup and restore
- Firmware updates
- Low Level Service Processor Resets
- Nodes and I/O module configuration and management
- Diagnostics
- Network management
- Chassis and compute node initialization

> **Note:** The IBM Flex System Enterprise Chassis supports up to 84 local CMM user accounts, and allows up to 32 simultaneous sessions.

Management Interfaces

The IBM Flex System Chassis Management Module allows two user interfaces to perform management functions:

- IBM Flex System Chassis Management Module web-based graphical user interface (GUI), that provides a way to perform CMM functions within a supported web browser. This user interface is a program that is part of the firmware on IBM Flex System Chassis Management Module.
- The IBM Flex System Chassis Management Module command-line interface (CLI) provides direct access to IBM Flex System management functions as an alternative to using the web-based user interface.

CMM CLI (SSH/Telnet), Web and SNMP management interface is accessible remotely over a Ethernet network. No crossover cable is required to connect to CMM if a local PC is used, as CMM performs automatic media dependent interface crossover

Web-based Graphical User Interface (GUI)

The IBM Flex System Chassis Management Module web interface communicates with the management program to perform chassis management tasks, such as:

- Controlling the chassis, compute nodes, and other components.
- Define login IDs and passwords
- Discover other chassis on the network and enabling access to them
- Access and configure I/O modules.
- Monitor the status of the compute nodes and other components
- Select recipients for alert notification of specific events
- Using a remote console for the compute nodes
- Change the startup sequence in a compute node
- Configure security settings such as data encryption and user account security

For detailed overview of the IBM Flex System Chassis Management Module web-based user interface, visit
http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.cmm.doc/cmm_user_guide.html

Command-line interface (CLI)

Using the IBM Flex System Chassis Management Module command-line interface (CLI), you can issue commands to control the power and configuration of the CMM and other components that are in an IBM Flex System Enterprise Chassis.

The command-line interface also provides access to the text-console command prompt on each compute node through a Serial over LAN (SOL) connection.

> **Note:** You access the CMM CLI through a direct serial or ethernet connection to the CMM, through a Telnet connection to the IP address of the CMM or through a Secure Shell (SSH) connection to the CMM.

You can find detailed reference information about using the command-line interface at,
http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.cmm.doc/cmm_cli_guide.htm

3.2.2 Accessing the CMM

After initial connection, you must configure the IBM Flex System Enterprise Chassis and its components for your operating environment. When the chassis is started, all installed components will be discovered and the vital product data (VPD) of each component is stored in the IBM Flex System Chassis Management Module.

The CMM will automatically configure the remote management port of the CMM so that you can configure and manage all of the chassis components. By default, the CMM is configured to respond to Dynamic Host Configuration Protocol (DHCP) first before using its static IPv4 address. It has the following default settings:

IP address 192.168.70.100
Subnet 255.255.255.0
User ID USERID (all capital letters)
Password PASSW0RD (all capital letters, with a zero instead of the letter O)

Note: The client computer that you connect to the CMM must be configured to operate on the same subnet as the CMM. The IP address of the CMM must also be in the same local domain as the client computer.

The default information, MAC address, and IPv6 link-local address is available in the network access card attached to all new CMMs, as shown in Figure 3-4.

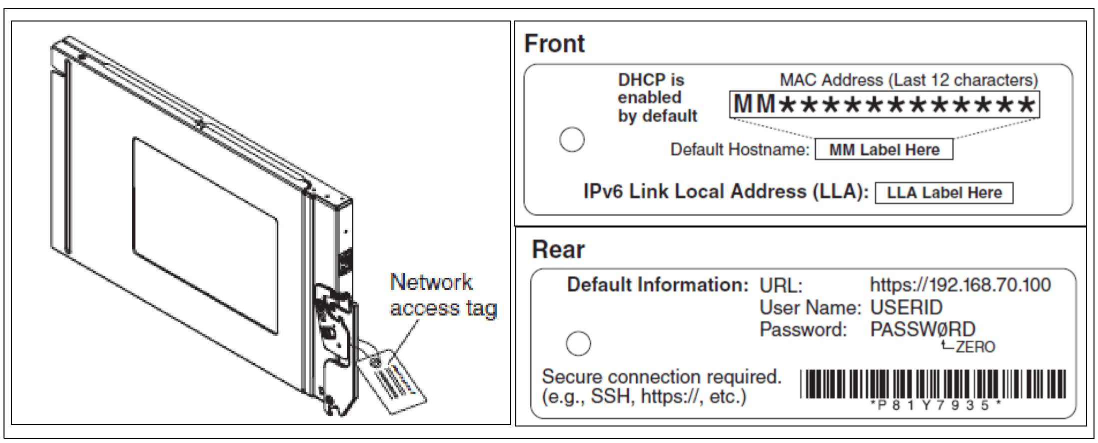


Figure 3-4 Network Access Tag

Note: The default security setting is *secure*, so HTTPS is required to connect to the CMM.

Logging in to the CMM

The following section gives step by step instructions regarding logging in to the CMM.

1. Open a web browser, and type in the IP address of the CMM.
2. At the Login screen, enter your User name and Password, and set the Inactive session time-out value, as shown in Figure 3-5 on page 91.

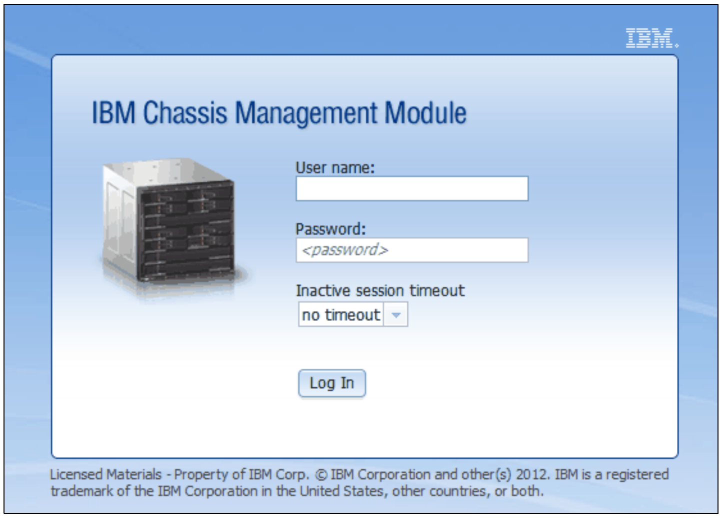


Figure 3-5 IBM Chassis Management Module Login screen

An Initial Setup wizard starts automatically the first time you access the web interface of a new CMM or a CMM that has been reset to its default settings, for more information regarding Initial setup of the CMM, refer to *Implementing Systems Management of IBM PureFlex System*, SG24-8060.

3. After a successful Login, CMM will open the default view showing Chassis Map, as shown in Figure 3-6.

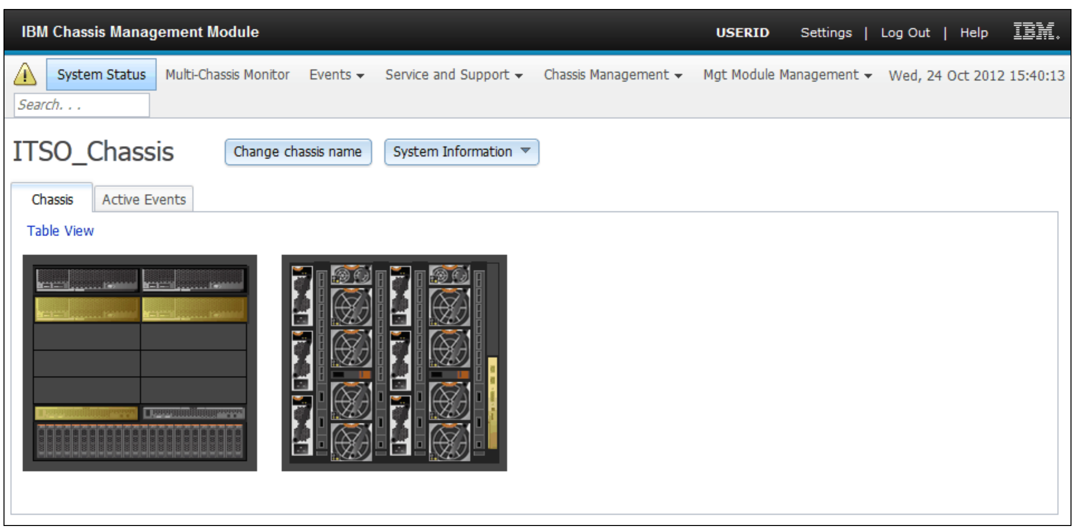


Figure 3-6 IBM Flex System Chassis Management Module chassis map (default) view

Disclaimer: In Figure 3-7 on page 92, and also in some preceding figures you may find a *yellow* exclamation point on the top left of the CMM interface along with yellow boxes around some components indicating a error condition. The reason of this visual discrepancy is that the various components of the chassis were being tested during the time of writing of this book.

3.2.3 Viewing and Configuring IP Addresses of Chassis Components

Due to the increasing demand for Remote Access to IT equipment over networks, being able to easily view and configure IP addresses of devices in a solution is of a great advantage to IT Administrators and Users. The IBM Flex System IBM Flex System Chassis Management Module gives options to conveniently find the IP information of all chassis components in a single view, and gives the ability to easily configure the networking options for a device.

- In IBM Flex System Chassis Management Module web-interface, a dedicated option is given in the top-level menu item: **Chassis Management** → **Component IP information** as shown in Figure 3-7.

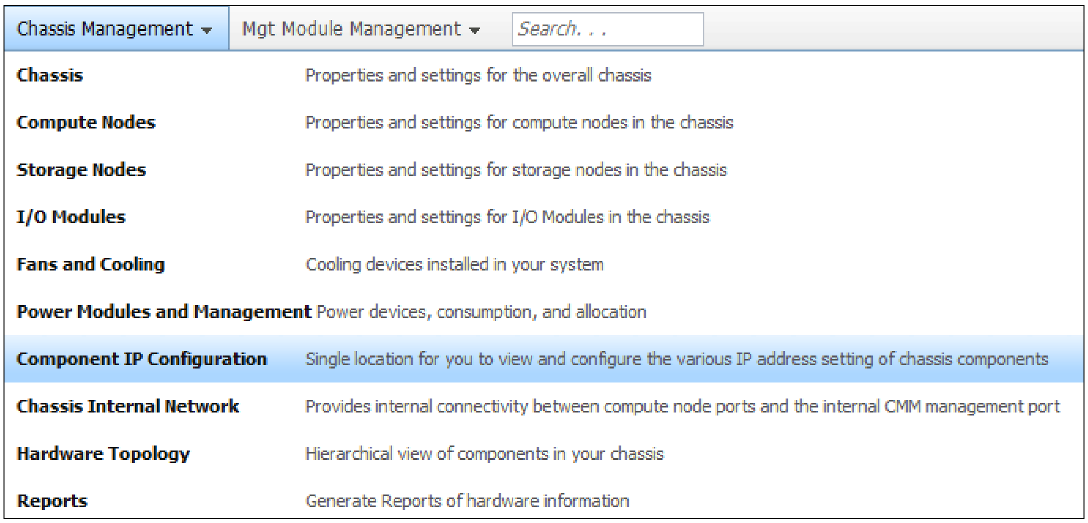


Figure 3-7 Component IP Configuration

- A window opens, giving a single location to view the IP addresses of various chassis Components. If you want to view IP address of a certain component instantly, you can just scroll the mouse pointer over View option, it opens a pop-up window that shows the IPv4 and IPv6 addresses of that device, as shown in Figure 3-8 on page 93.

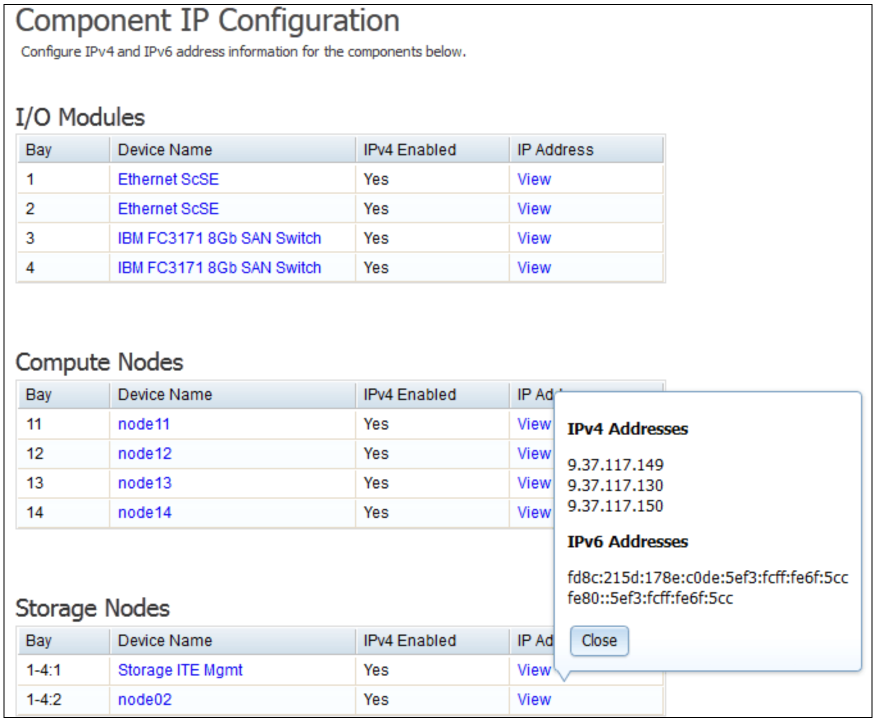


Figure 3-8 Component IP Configuration view

- To configure the IP address of a certain chassis component, click on the component (*under Device Name*), as shown in Figure 3-9.

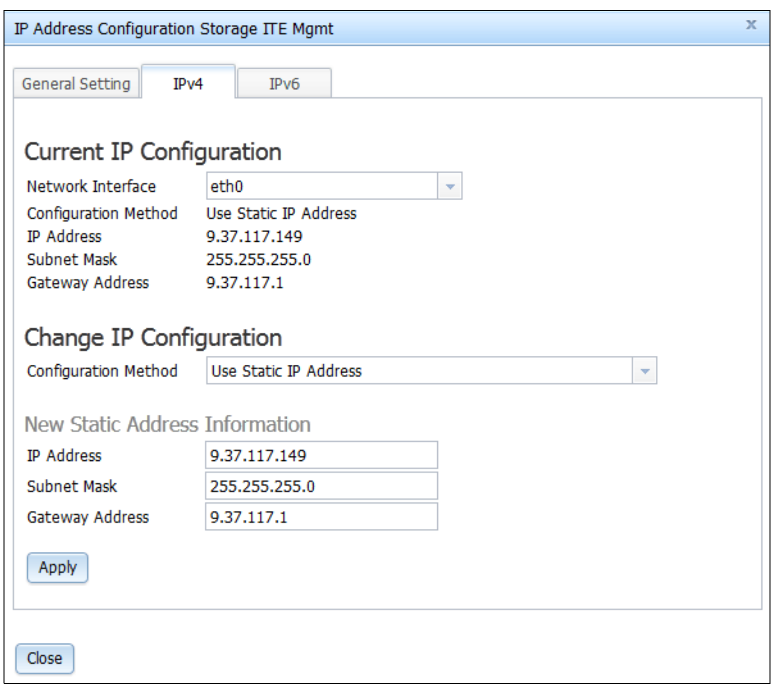


Figure 3-9 IP Address Configuration

You can also view and configure IP addresses using IBM Flex System Manager as well as the node-specific management controllers.

There are numerous functions and features offered by IBM Flex System Chassis Management Module that are beyond the scope of this publication, however, we will cover the I/O module access features in the next section, due to its relevance with Storage management.

For detailed information regarding IBM Flex System Chassis Management Module configuration capabilities, visit
http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.cmm.doc/cmm_ui_config_the_cmm.html

or refer to *Implementing Systems Management of IBM PureFlex System*, SG24-8060.

3.2.4 Accessing I/O Modules using CMM

IBM IBM Flex System Chassis Management Module performs end-to-end management of all components in the solution. It allows many functions and features that can be performed on I/O Modules, as well as, it gives access to platform-based management interfaces of the modules for in-depth management.

To access the properties and configuration of I/O Modules through IBM Flex System Manager, click on ***Chassis Management*** → ***I/O Modules***, as shown in Figure 3-10.

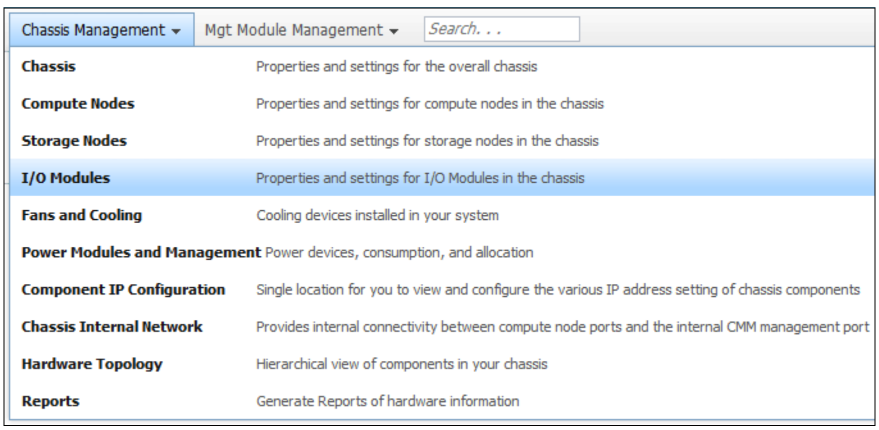


Figure 3-10 FSM - Accessing I/O Modules

The next view shows the available I/O modules in IBM Flex System. It also gives options to Power and Restart the modules as well as the actions that can be performed on them, as shown in Figure 3-11.

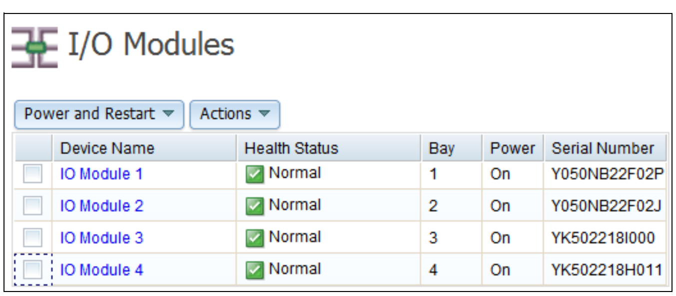


Figure 3-11 Viewing I/O Modules in FSM

Clicking on an IO Module (*Device Name*) opens up a window showing its properties, as shown in Figure 3-12.

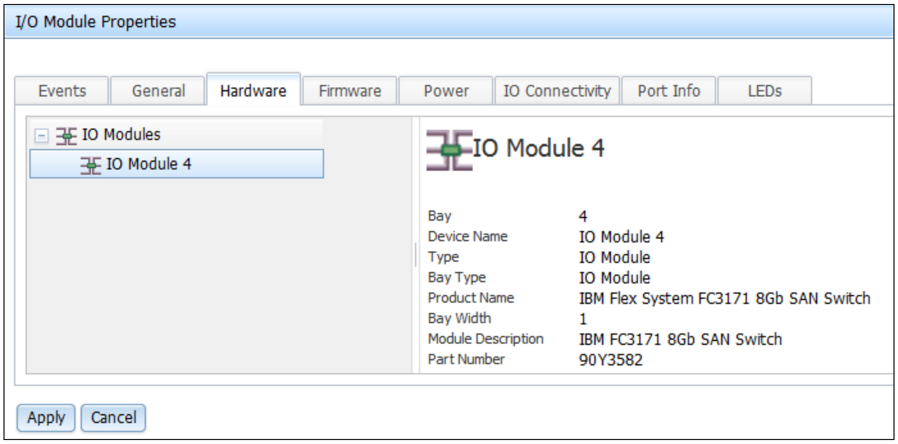


Figure 3-12 Properties of IO Module device

The Properties window has multiple tabs, showing detailed information regarding the module. Click on the *IO Connectivity* tab to view the connectivity details between the IO module and the compute nodes in the IBM Flex System. This is shown in Figure 3-13.

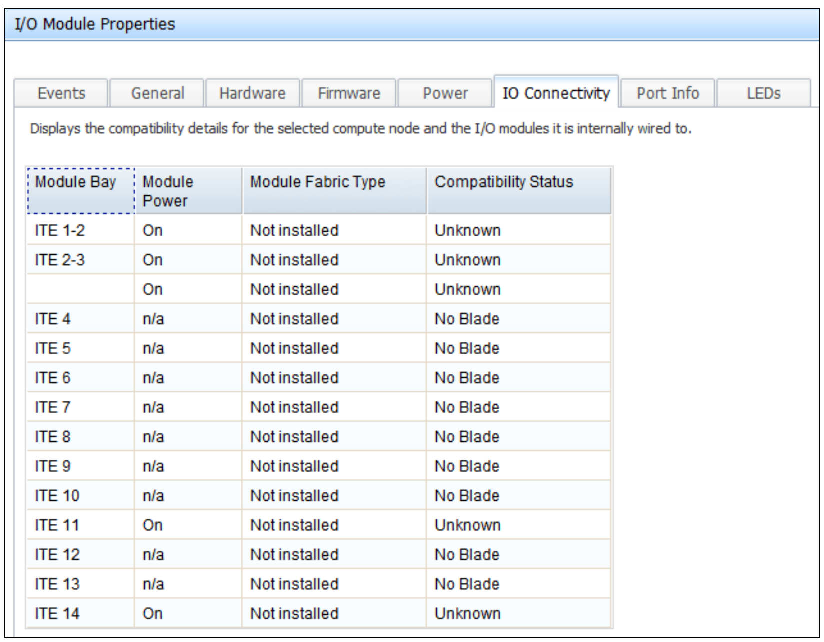


Figure 3-13 IO Connectivity Status of an IO Module

To open the Module-specific Management interface, click on the desired *IO Module* then *Actions* drop-down button, as shown in Figure on page 96.

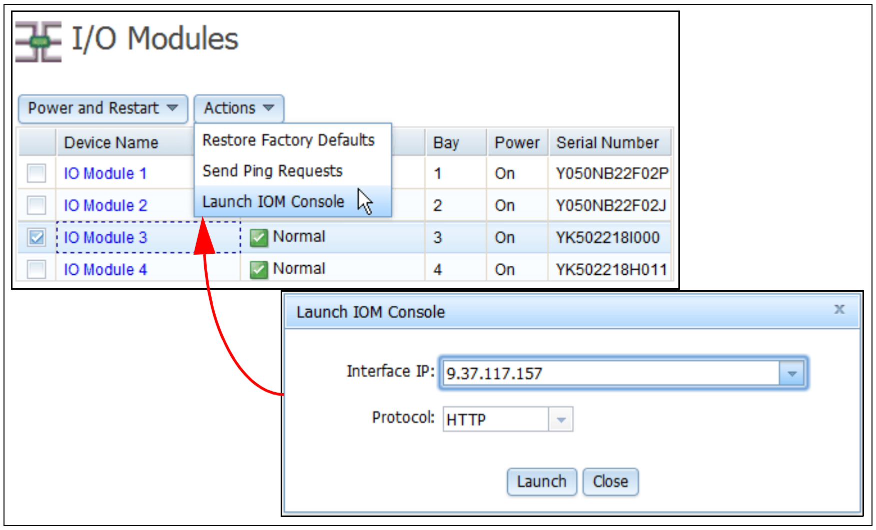


Figure 3-14 Launching IOM Console via CMM

The drop-down menu gives the option to select HTTP or HTTPS protocol. In case of most IO Modules, HTTPS may need to be selected. This will result in execution of Switch Modules platform-level management interface.

3.2.5 Managing Storage using IBM Flex System Chassis Management Module

After completing initial setup, a subsequent login into CMM opens the *System Status* page. The top of the page identifies the current user and contains a *Settings* menu and *Log Out* option, as shown in Figure 3-15.

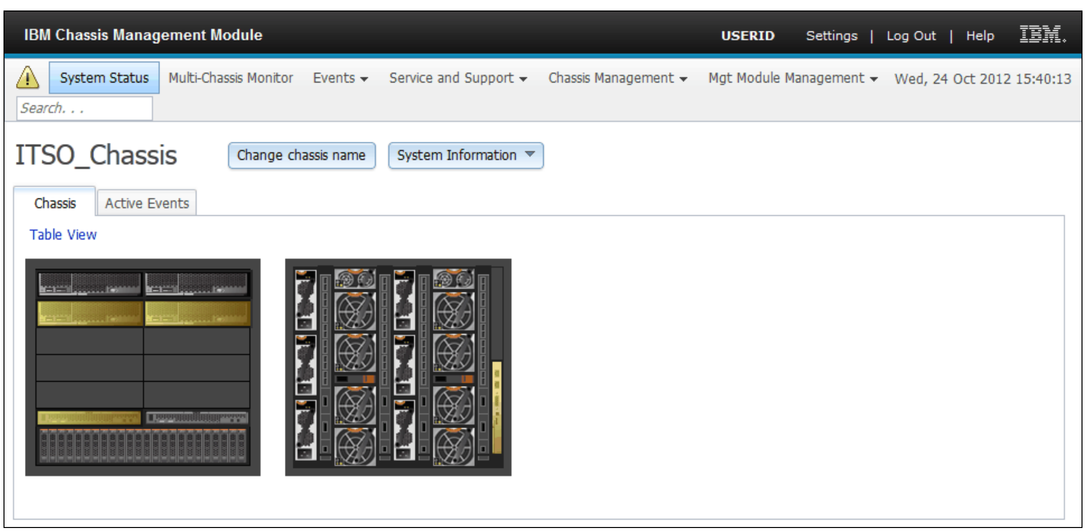


Figure 3-15 IBM Chassis Management Module chassis map

The *System Status* shows the view of Chassis, front and back. It also shows health status of the individual physical components.

> **Note:** Except for the top menu items, the remaining content of the *System Status* page and other web interface pages will vary according to the type of IBM Flex System Enterprise Chassis that you are using and the firmware versions and options that are installed.

Scrolling mouse on the selected component opens a pop-up windows that lists its name, location and FRU number, as shown in Figure 3-16.

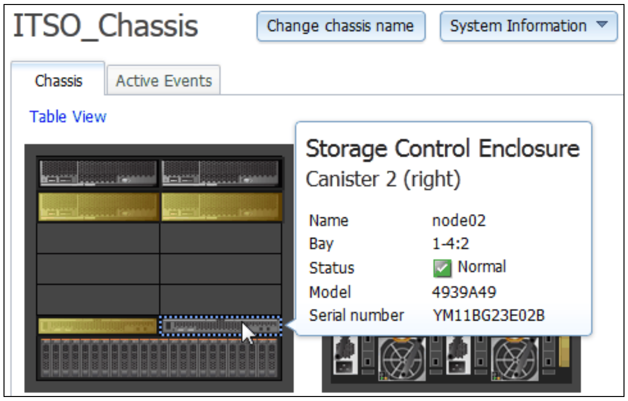


Figure 3-16 CMM chassis view - component pop-up

When you click on a component, details about it emerge in the lower part of interface. Depending on the component, one or more tabs show information regarding hardware, firmware, events, and other information. An example is shown in Figure 3-17.

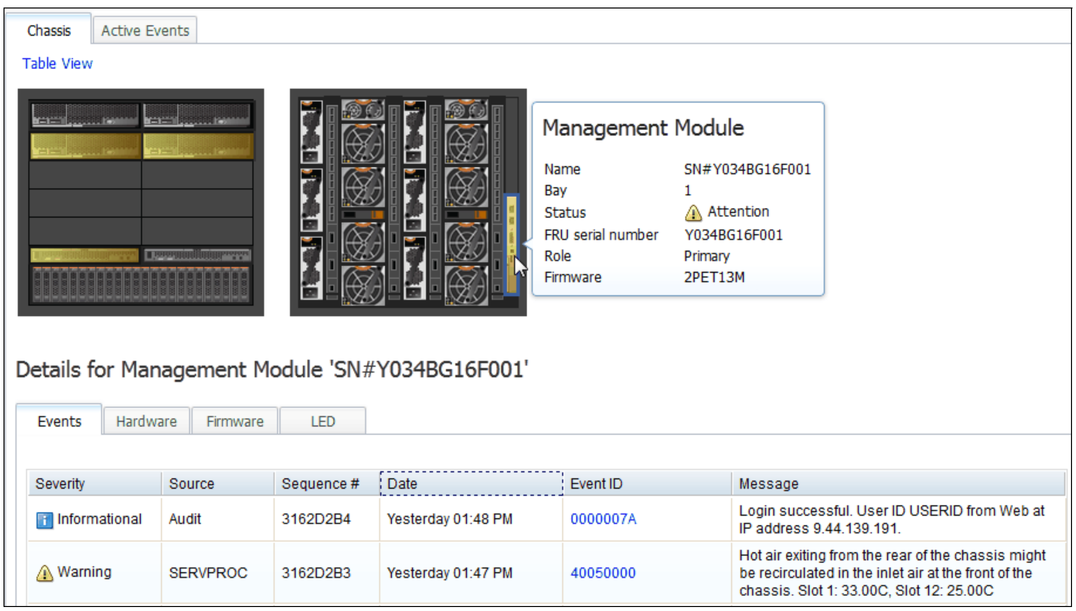


Figure 3-17 CMM - Details of the selected Chassis component

The *System Information* (blue) button which is located above the graphical representation of the Chassis, gives a quick view of the selected component's information. The quick view includes part information, serial number, health and temperature, and unique identity (depending on the component), as shown in Figure 3-18 on page 98.

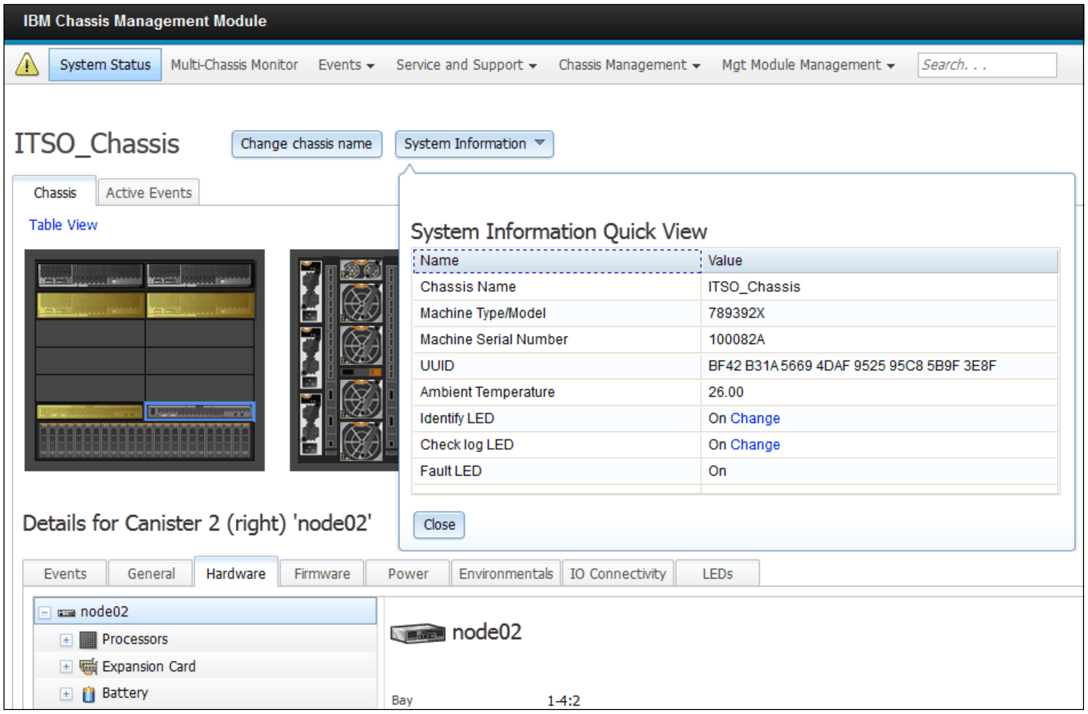


Figure 3-18 CMM - System Information Quick View

Selecting a component also opens up an Actions frame, which lists the actions that can be performed on that particular IBM Flex component, as shown in Figure 3-19.

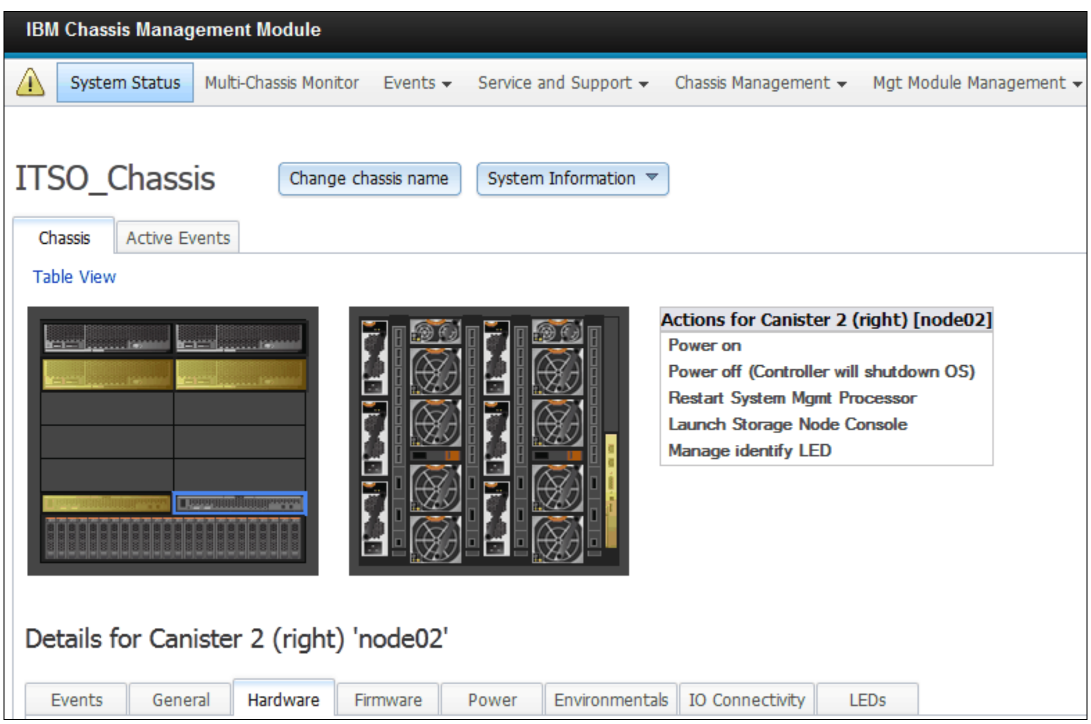


Figure 3-19 CMM - Actions for Chassis Component

In the Figure 3-19, an IBM Flex System V7000 Storage Node canister is seen as selected (blue rectangle around component). On the right-side of the Chassis rear view, there is a list

of Actions for the particular Canister, which includes options to power on/off, restart the service processor, launch the detailed IBM Flex System V7000 Management application, as well as manage the identity LEDs on the component.

The area of GUI under the graphical representation of the chassis shows details regarding the selected component. This includes very detailed information laid out in convenient tabular format. The example in Figure 3-20 shows firmware information including its type, build info, release date, revision and role.

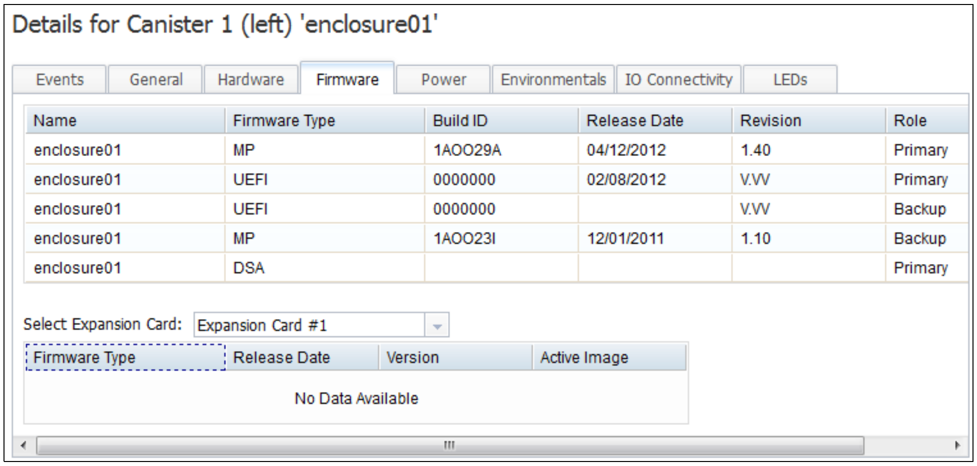


Figure 3-20 Details regarding selected chassis component - Firmware tab

With the advent of complex integrated IT solutions, sophisticated hardware with varying power and voltage requirements get consolidated into a typical enterprise rack. To ensure proper functioning of all the hardware, IT administrators have to ensure that all equipment stays below the temperature, power and voltage thresholds at all times. The IBM Flex System Chassis Management Module web interface provides multiple views to facilitate monitoring and notification of environmental and temporal indices.

Figure 3-21 shows the Environmental details of various components within selected chassis hardware (storage node canister in this case). It shows the current values as well as the thresholds set for these components.

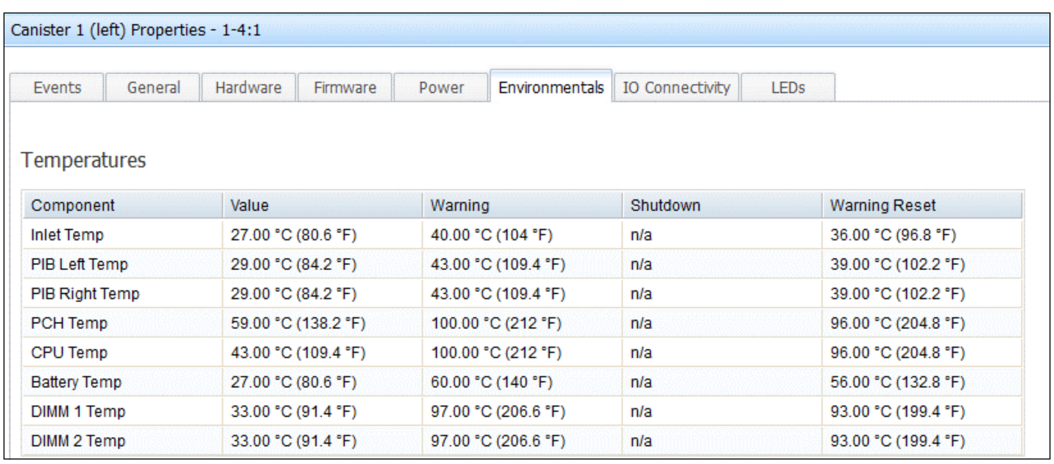
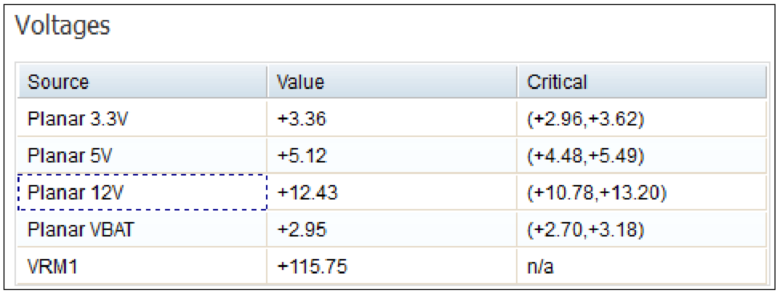


Figure 3-21 Canister properties - Environmental

Environmental view also reports voltage value and thresholds, as shown in Figure 3-22 on page 100.

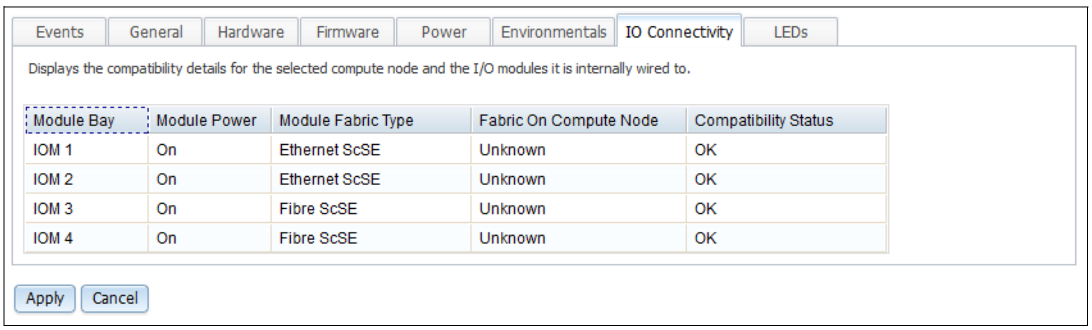


Voltages

Source	Value	Critical
Planar 3.3V	+3.36	(+2.96,+3.62)
Planar 5V	+5.12	(+4.48,+5.49)
Planar 12V	+12.43	(+10.78,+13.20)
Planar VBAT	+2.95	(+2.70,+3.18)
VRM1	+115.75	n/a

Figure 3-22 Canister properties - Voltages

An important view for storage canister is the IO Connectivity tab, which shows the fabric information related to the IO modules in the chassis, as shown in Figure 3-23.



Module Bay	Module Power	Module Fabric Type	Fabric On Compute Node	Compatibility Status
IOM 1	On	Ethernet ScSE	Unknown	OK
IOM 2	On	Ethernet ScSE	Unknown	OK
IOM 3	On	Fibre ScSE	Unknown	OK
IOM 4	On	Fibre ScSE	Unknown	OK

Figure 3-23 IO Connectivity

Figure 3-24 on page 101 shows the DC Power consumption for the component.

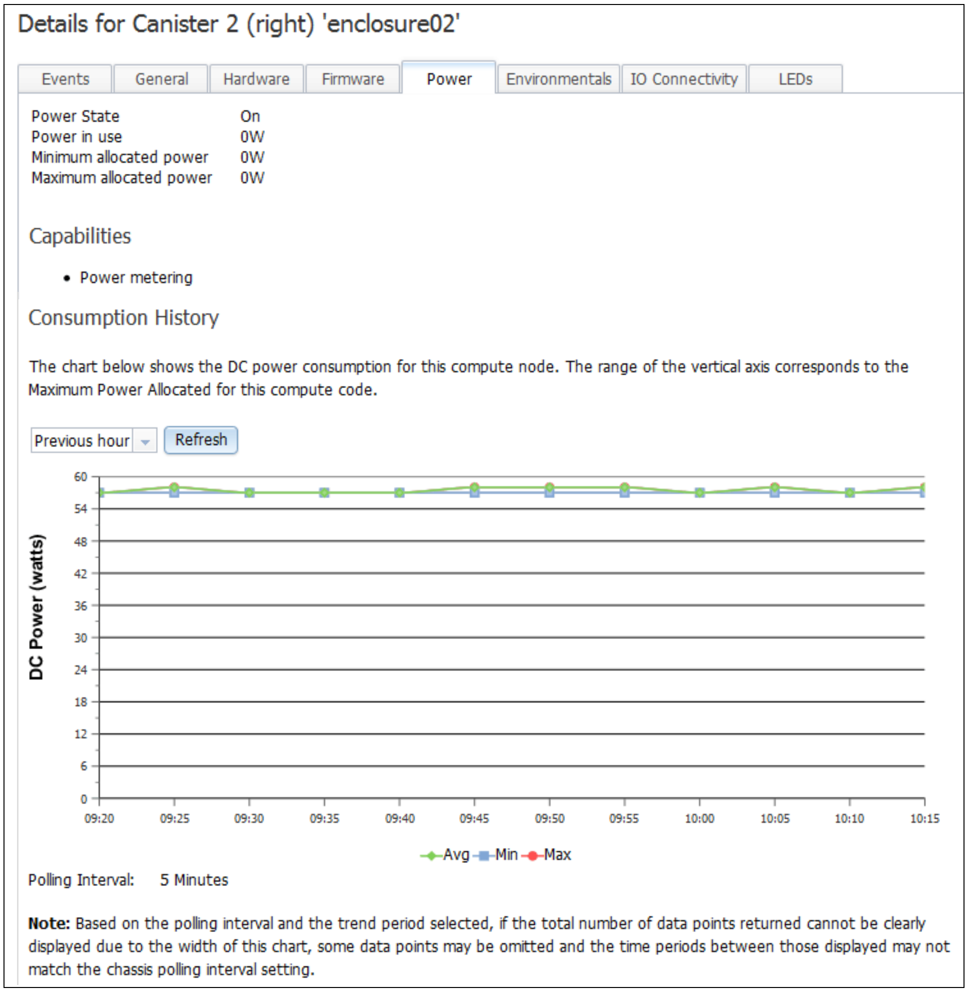


Figure 3-24 Power consumption of selected chassis component

The IBM Flex System Chassis Management Module web interface has a menu structure at the top of each page that gives easy access to most management functions. One of the most frequently used menu item is Chassis Management, which shows properties, setting and various views and reports regarding the nodes, modules and other components in the Chassis.

Figure 3-25 shows the option that gives the hierarchical view of chassis components.

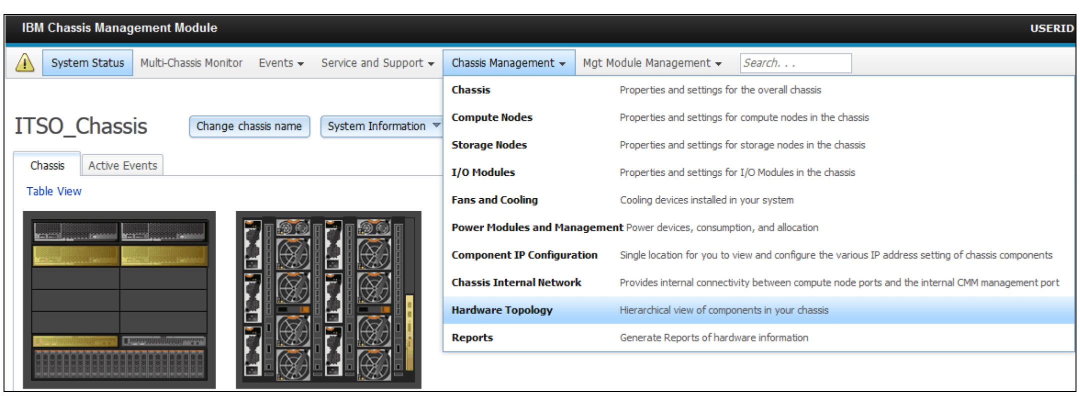


Figure 3-25 Hardware Topology view from Chassis Management menu

The Chassis Hardware Topology view provides very detailed information regarding all the hardware in the chassis. It goes deep to the point where you can select individual drives, batteries and even expansion cards within a storage or compute node, as shown in Figure 3-26.

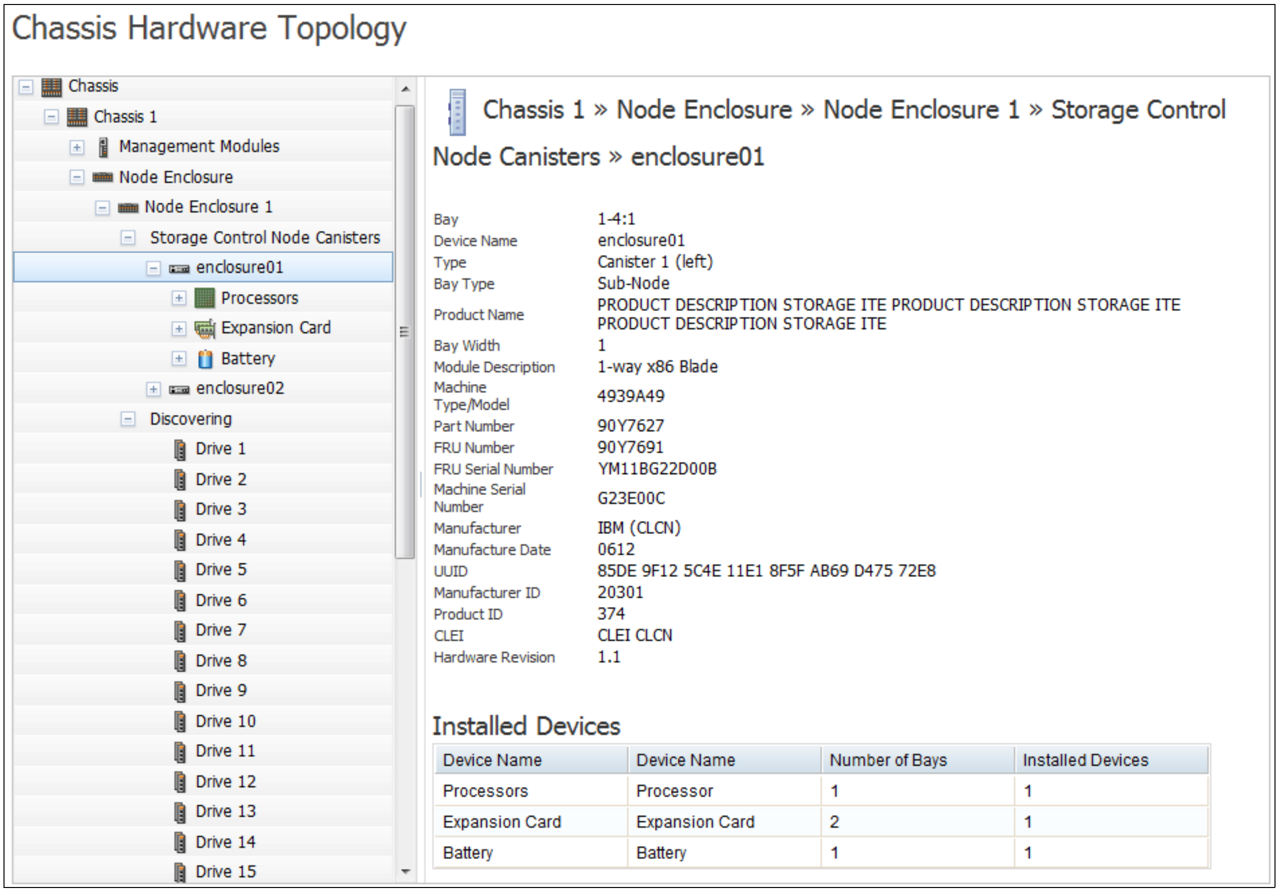


Figure 3-26 Chassis Hardware topology

- From the top menu, select **Chassis Management** → **Storage Nodes** to view details regarding storage nodes, as shown in Figure 3-27.

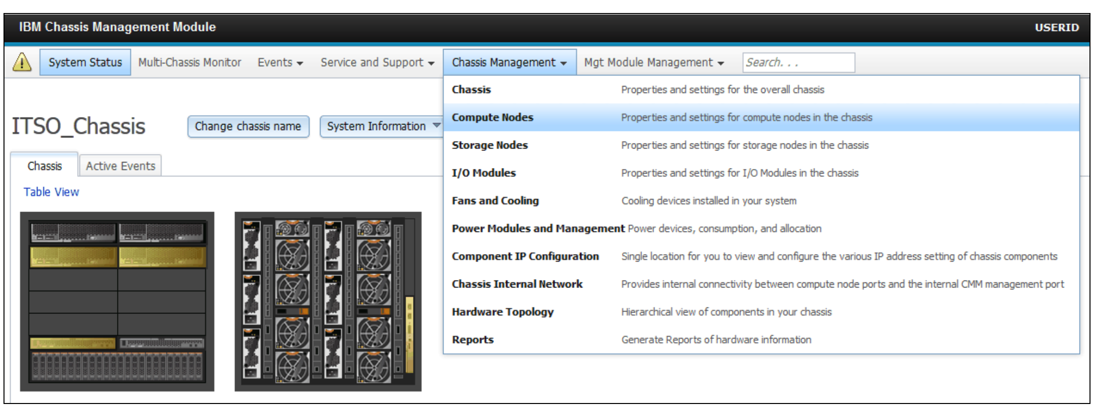


Figure 3-27 Chassis Management - Storage Nodes

The Storage Nodes view shows all the storage enclosures in the chassis, and gives options to perform various actions on these, including power and restart options, as shown in Figure 3-28.

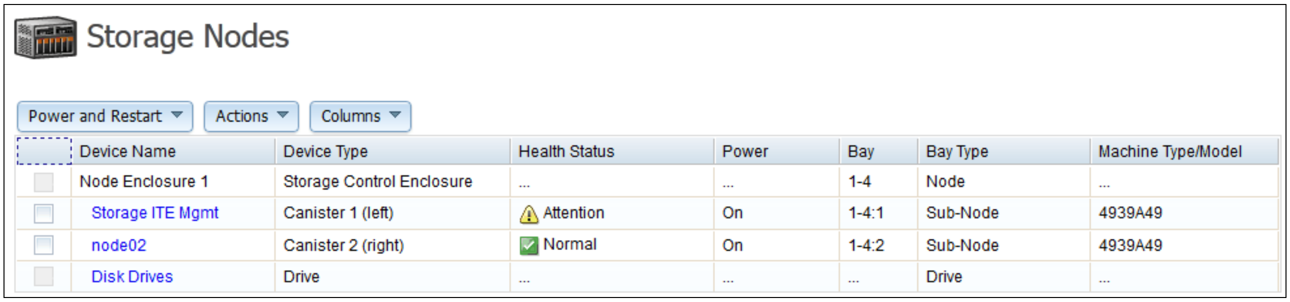


Figure 3-28 CMM - Storage Nodes - view of all enclosures

In the Storage Node view, you can select the column headers depending on how much detail is desired, as shown in Figure 3-29.

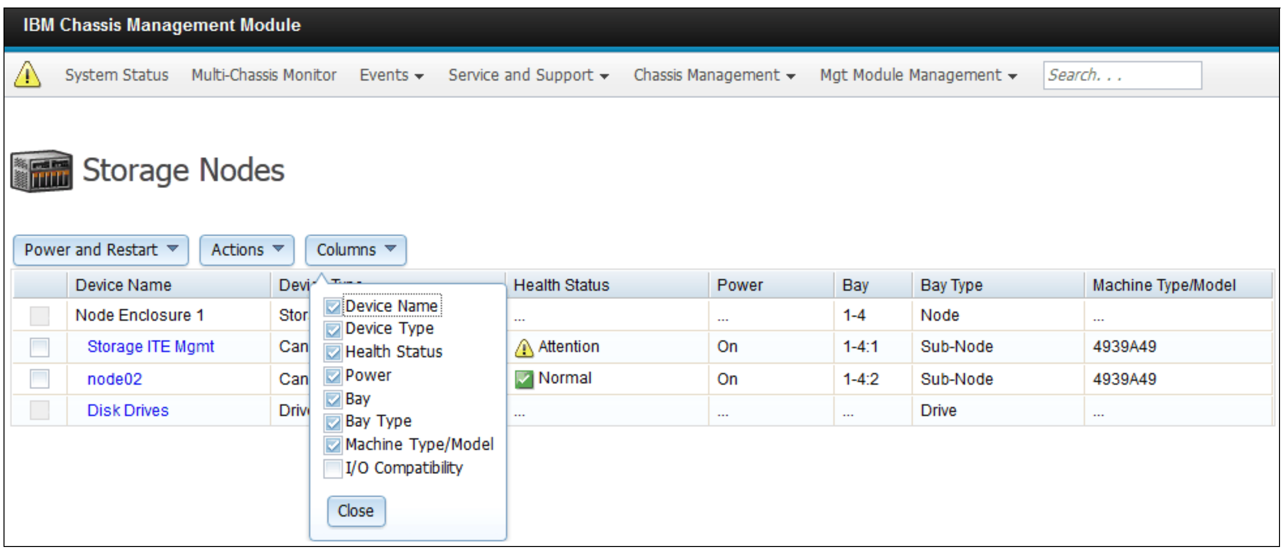


Figure 3-29 CMM - Storage Nodes view - columns

Actions menu gives options to manage identify LED of the selected enclosure, as shown in Figure 3-30 on page 104.

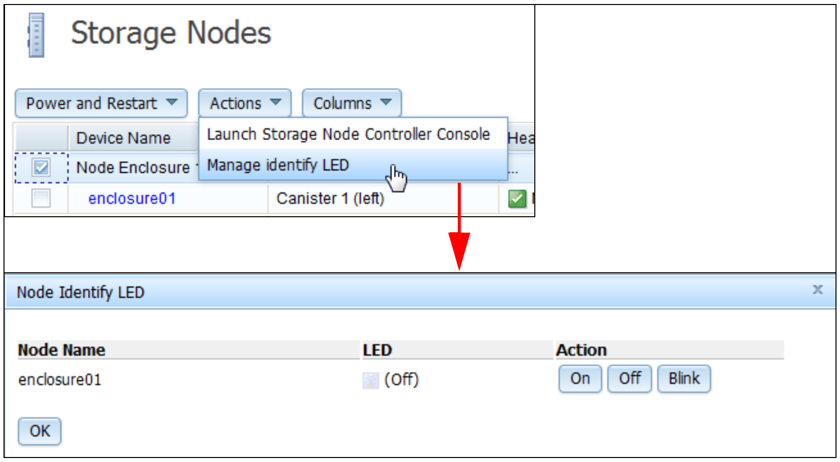


Figure 3-30 Manage Identify LED

The IBM Flex System V7000 Storage Node User Interface is used for detailed-level management and configuration of the IBM Flex System V7000 Storage Node.

Click on **Actions** → **Launch Storage Node Controller Console**, which will take you to the IBM Flex System V7000 Storage User Interface, as shown in Figure 3-31.

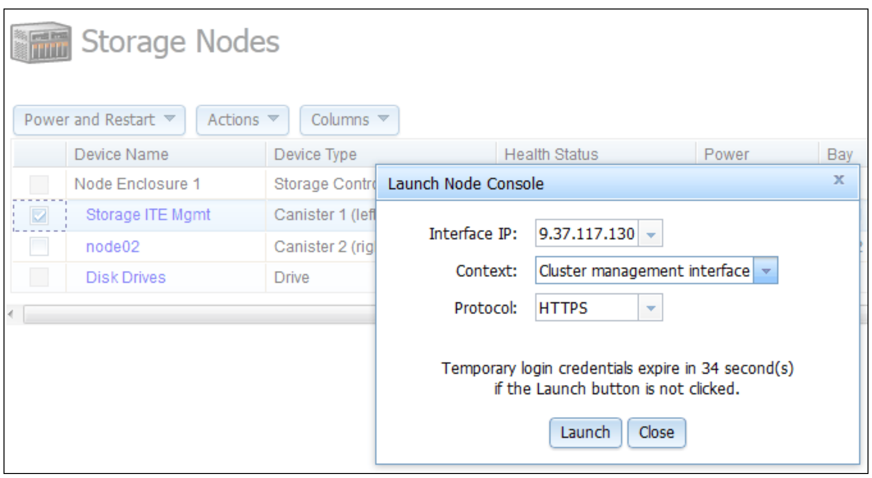


Figure 3-31 CMM - Launch Storage Node Controller Console

Next window is the IBM Flex System V7000 Storage Management application's login screen, as shown in Figure 3-32 on page 105.

Figure 3-32 IIBM Flex System V7000 Storage Node Login window

After logging in, you can use the IBM Flex System V7000 Storage User Interface directly to perform the management tasks, as shown in Figure 3-33.

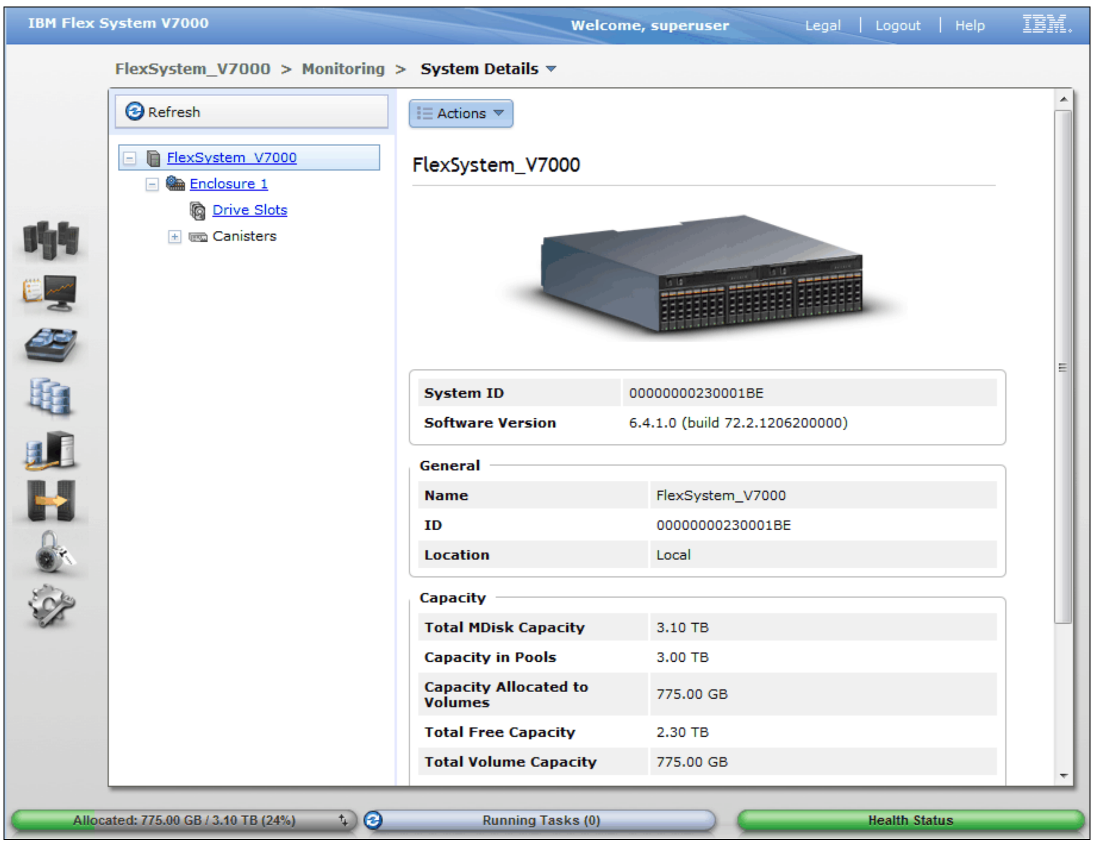


Figure 3-33 IBM Flex System V7000 System Details

For further information regarding IBM Flex System V7000 Storage User Interface, refer to Chapter 5, “IBM Flex System V7000 Storage Node Graphical user interface overview” on page 175.

3.2.6 Data collection using CMM

IBM Flex System Chassis Management Module offers various powerful options to pull service information from the components in chassis. The *Service and Support* menu allows the user to view list of serviceable problems and their current status, enable and configure electronic Call Home capability, review detailed status about CMM's health and connectivity, and generate as well as retrieve service data files from the chassis components.

For acquiring service data from the Storage Node, you can perform the following steps:

- From the top menu, select **Service and Support** → **Download Service Data**, as shown in Figure 3-34.

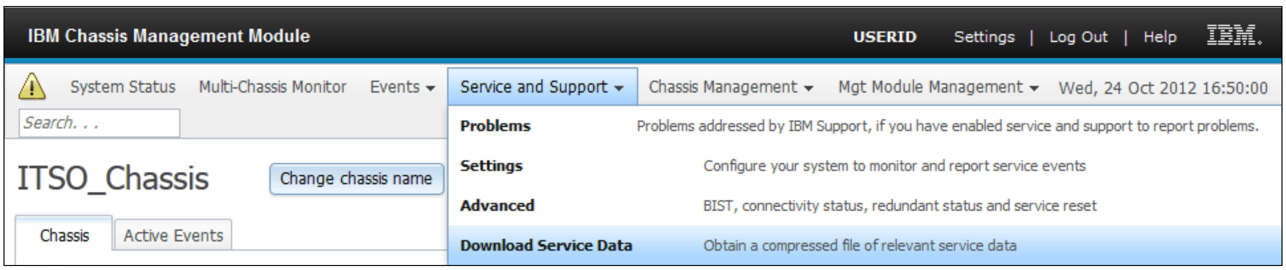


Figure 3-34 CMM - Service and Support - Download Service Data

- The *Download Service Data* page gives four options. We pick the *Storage Node Service Data* as shown in Figure 3-35.

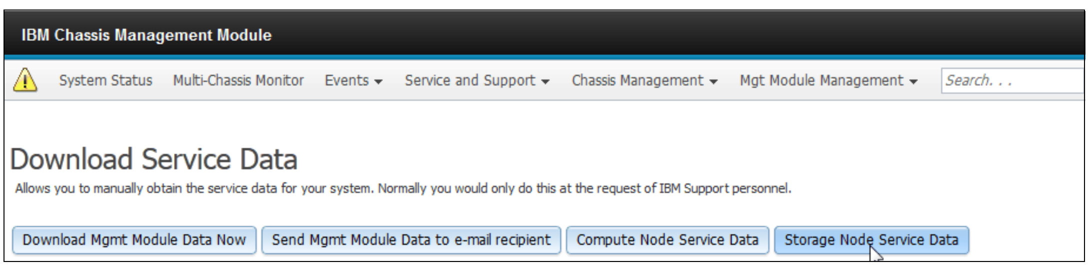


Figure 3-35 CMM - Download Service Data from Storage Node

This next window gives three options to dump the Storage Node Service Data. Description of all three is also provided, so user can easily choose the desired option.

- In our case, we click on the drop-down menu *Dump Type:* then select *Service Data*. Choose the option *Initiate and Collect*, which allows storage node to initiate the generation of a new dump file on the compute node and overwrite any existing file. The CMM will then collect the dump file by transferring it from the node to the CMM file system. This is shown in Figure 3-36 on page 107.

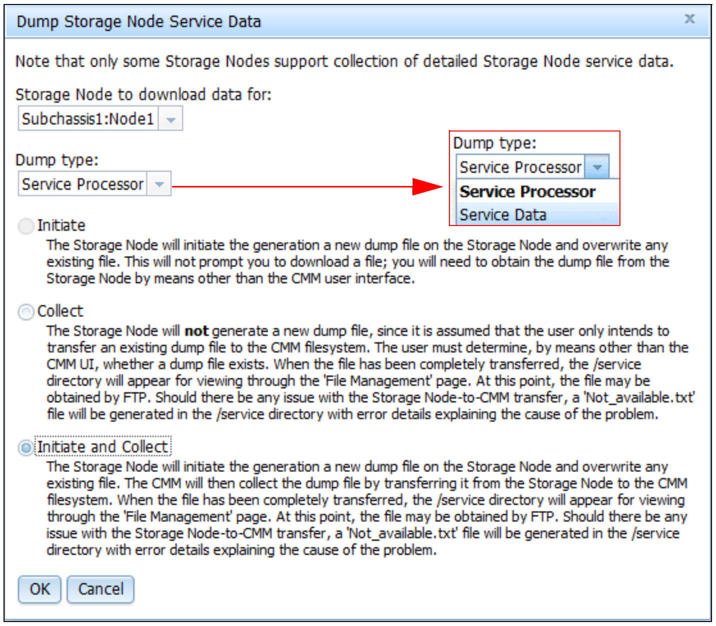


Figure 3-36 Initiate and Collect

- System will prompt you for verification, indicate that operation is in progress, and then its completion, as shown in Figure 3-37.

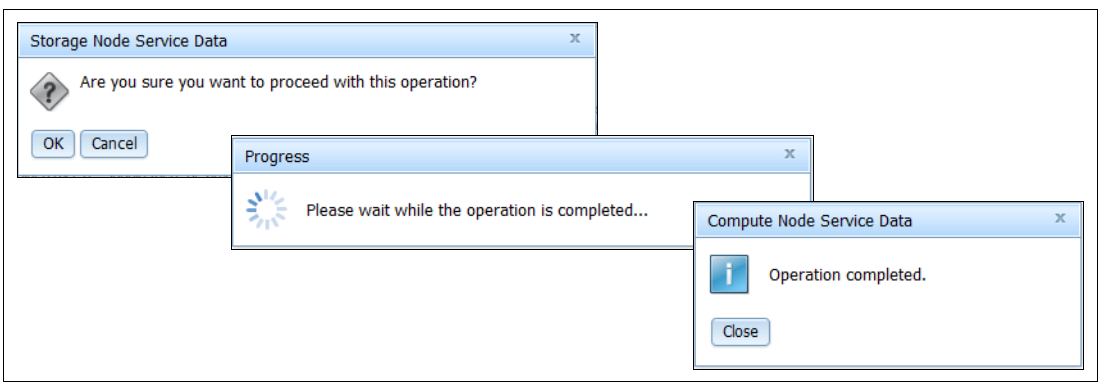


Figure 3-37 Status dialog boxes

To view or delete files in the CMM local storage file system, select **Mgt. Module Management** → **File Management**, as shown in Figure 3-38.

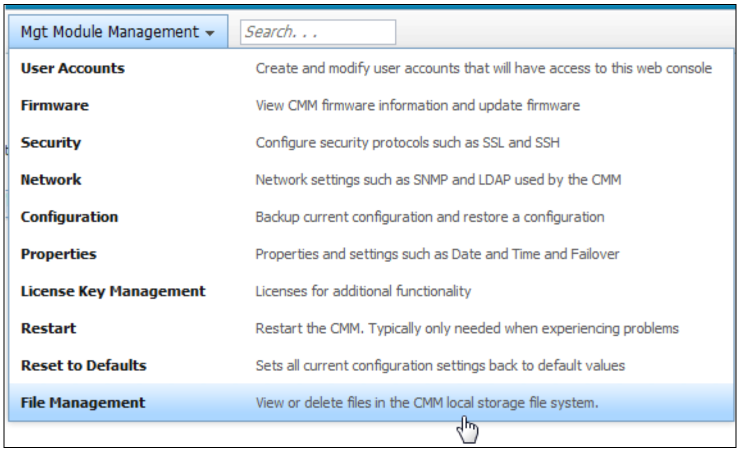


Figure 3-38 File Management

File Management page shows the status of CMM's local Storage file system and the CMM File management menu as shown in Figure 3-39.

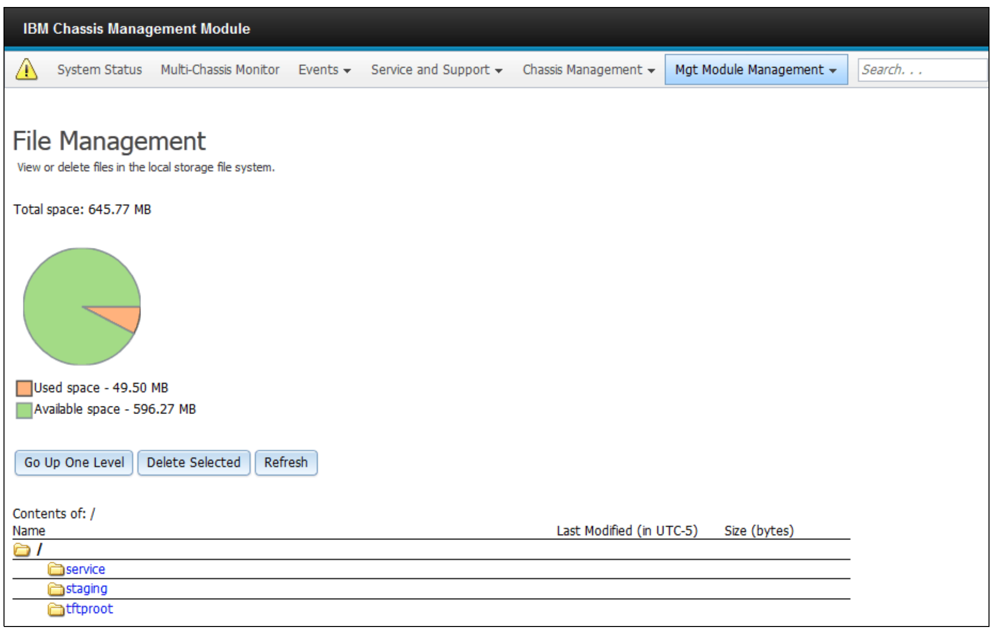


Figure 3-39 Management of local CMM file system

/service/ folder contains the file that contains the data dump, as shown in Figure 3-40 on page 109.

Figure 3-40 Viewing service data in /service/ folder

The file containing dump of service data is also copied in /tftproot/service/ for ease of transfer, as shown in Figure 3-41.

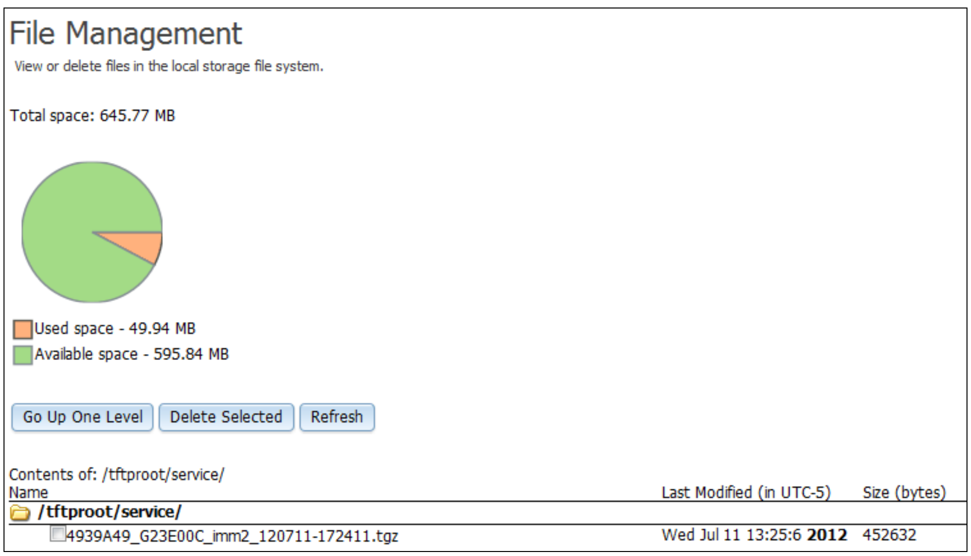


Figure 3-41 Viewing service data in /tftp/service/ path

For more details regarding the IBM IBM Flex System Chassis Management Module, visit `http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.cmm.doc/cmm_product_page.html`

3.3 Flex System Manager (FSM)

The IBM Flex System Manager Node (FSM) is an advanced chassis management appliance which provides the ability to manage up to four chassis. Also referred to as the Management Node, FSM is based on an Intel-based x86 compute node that comes with preloaded management software. FSM software contains a fully operational virtual infrastructure management stack for managing multiple chassis from a single focus-point.

IBM Flex System Manager provides an instant and resource-oriented integrated view of chassis and chassis resources, including compute and storage nodes as well as I/O modules. It simplifies management by facilitating faster implementation, reconfiguration and provisioning of server, storage and network resources. It acts as a single-point of support across all assets and resources in IBM PureFlex Systems for the complete range of management tasks, including service requests, incident handling and problem resolution, as well as handling specialized management tasks in connection with security, performance, resource allocation, and application version control, etc.

Figure 3-42 shows the front view of an IBM Flex System Manager Node:

Figure 3-42 IBM Flex System Manager Node

> **Note:** The IBM Flex System Chassis Management Module provides control and information on the storage hardware, while the IBM Flex System Manager in addition provides configuration functionality.

3.3.1 Overview of FSM

The comprehensive management capabilities of an IBM Flex System Manager allows for seamless and simplified storage management, virtualization and networking.

Consolidating management and administration tasks into a single, consistent interface while automating many tasks across various compute and storage nodes simplifies the job of administrators to manage and configure servers, storage, network and other resources.

Whole rack management of energy, power and cooling, etc. reduces cost and cuts downtime, while the systems operate at higher performance levels and are more consistent in operation. The single interface eliminates the need of having different administrative support and skills for each platform, thus reducing training costs.

The IBM Flex System supports a secure policy by default to ensure a secure chassis infrastructure, and a Legacy policy that offers flexibility in chassis security. All components run the same security policy. A centralized security policy makes Enterprise Chassis easy to configure. The system password is required by the management software to perform system-management functions, such as enabling access to the local user registry. The

management software system password is specified when the management software is set up initially, during the Management Server Setup wizard. For more information regarding FSM initial setup, refer to *Implementing Systems Management of IBM PureFlex System*, SG24-8060.

IBM Flex System Manager will auto-discover new and existing components, as well as automatically initiate and perform set-up tasks. As part of this discovery process, IBM Flex System Manager prepares and maintains complete, detailed front, back and chassis map views of each chassis and all physical compute nodes and virtual servers that exist in the rack.

Management and control of the physical and virtualized infrastructure includes initiating the detailed tasks required for full life-cycle management of virtualized devices, including their creation, provisioning, configuration, operation and de-provisioning. It can build and maintain a library of pre-defined virtual devices. This allows fast automated, on-demand adjustments in server, storage and network capacities in response to changing business and operational needs.

The built-in network provides a reliable and secure communications path for management and inter-rack data sharing. Sophisticated network resource management moves the focus from managing the network to managing the workload. The complete infrastructure is managed to assure maximum workload performance. This allows for operational environment-aware automated smart placement and provisioning of the workload itself. The emphasis is on optimized, automated, integrated network control and management.

Other significant characteristics of an IBM Flex System Manager are:

- All basic and advanced functions come preloaded as an appliance.
- A real-time, multi chassis view of hardware components is presented with overlays for additional information.
- Easy-to-use multi-chassis management capabilities include;
  - Quick Start Wizards with automated discovery
  - Advanced remote presence console across multiple chassis
  - Centralized IBM Feature on Demand (FoD) license key management
- Integrated x86 & Power nodes, Storage & Network management
- Virtualization Management including resource pools
- Robust Security (Centralized User Management, Security Policy, Certs)
- Integrated LDAP and NTP servers for private management network
- Active Energy Management
- Active/Passive Failover to redundant FSM Appliance
- Additional advanced management capabilities available through add-ons
- Upward integration into Systems Director, Tivoli and other 3rd Party Enterprise Managers
- Configuration Manager plug-ins (besides IBM Flex System Chassis Management Module)
  - Configure Chassis and Internal Switch Components
  - Configuration of Compute Nodes
- Remote Presence
  - Login and take over Compute node from Administrators console
  - Mount DVD drive remotely
  - Perform Bare Metal Install remotely

- Service & Support Management with Call-Home
  - Automatically detect and log HW Problems
  - Notify IBM directly of HW Outages
- Required update of the default passwords after the initial setup
- Only secure communication protocols, such as SSH and SSL. Unencrypted protocols, such as HTTP, Telnet, and SNMPv1, are disabled.
- Certificates to establish secure, trusted connections for applications that run on the management processors
- Strong password policies with automatic validation and verification checks

For further details regarding IBM Flex System Manager features and capabilities, refer to: *Implementing Systems Management of IBM PureFlex System*, SG24-8060

3.3.2 IBM Flex System Manager's storage management features

IBM PureFlex offers interoperability with a wide range of storage subsystems. The system infrastructure has the following salient features:

- IBM Flex System V7000 Storage Node - the Integrated virtualized IBM Flex System Storage
- IBM Flex System Storage Virtualization - to virtualize external Storage for greater data center efficiency and utilization (e.g. IBM Storwize V7000, IBM Flex System V7000 Storage Node, & IBM SAN Volume Controller,etc.)
- IBM Flex System Storage Interoperability - a broad set of IBM storage is supported with IBM Flex System. Interoperability with 3rd party storage is also offered via virtualization
- IBM Flex System FSM Storage Control, which offers the following:
  - Discovery and Inventory
  - Monitoring and Alerts
  - Configuration
  - Provisioning
- IBM Flex System FSM Advanced, which offers the following features and functions:
  - Integrated virtualization management across server, storage, network
  - Image repository and management Storage provisioning for image creation, deployment, and cloning
  - System Pools
  - Integrated management of storage in life cycle of defining and managing system pools
  - Virtual Image Cloning
  - Integrated storage provisioning and virtual image placement for new virtual machines

The interoperability is illustrated in Figure 3-43.

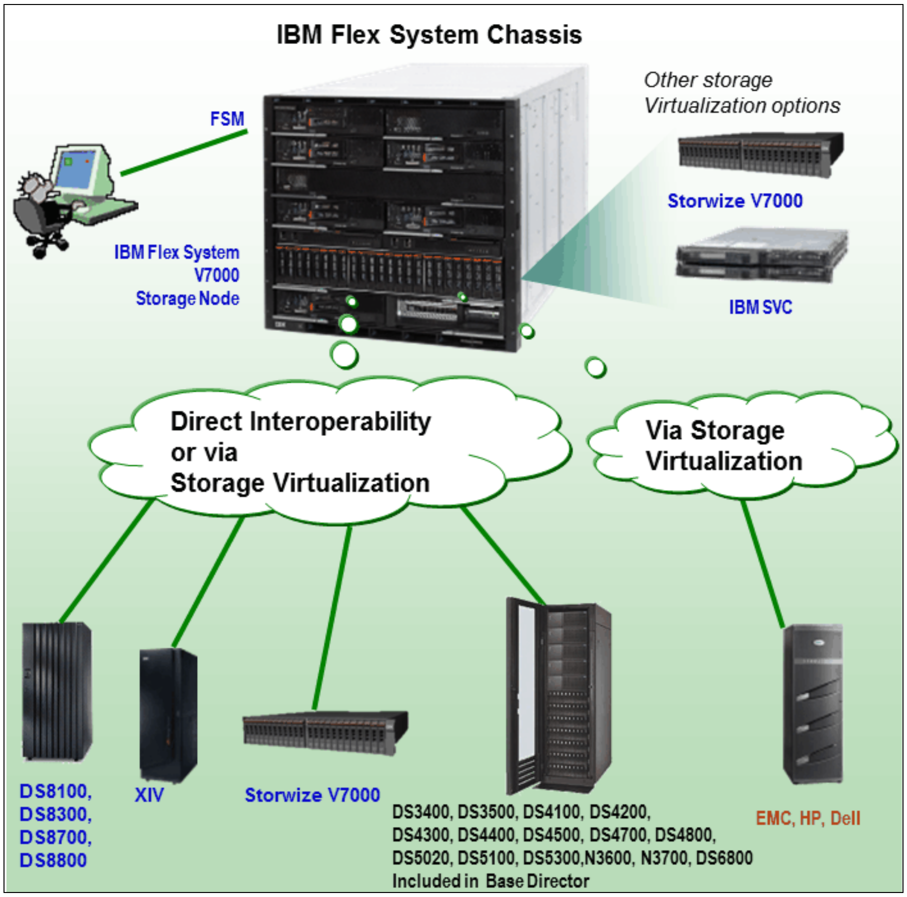


Figure 3-43 Storage Interoperability

IBM Flex System Manager is the common management interface used in an IBM Flex System for multi-chassis management. It offers features for simplified administration of storage along with server and network management. The integrated storage management with Flex System Manager offers simplified storage virtualization capabilities.

Storage Features in IBM Flex System Manager include:

– Storage device discovery
– Showing relationships between storage and server resources
– Physical and logical topologies
– Configuration of physical and logical storage
– Viewing controller and volume statuses
– Setting notification alerts
– Integration with VM Control provides:
  - Storage provisioning for virtual server image creation, deployment, and cloning
  - Storage system pool lifecycle management, provisioning, and cloning actions within the pool

You can virtualize and manage the complete solution infrastructure using the IBM Flex System Manager. Figure 3-44 shows FSM virtualization overview.

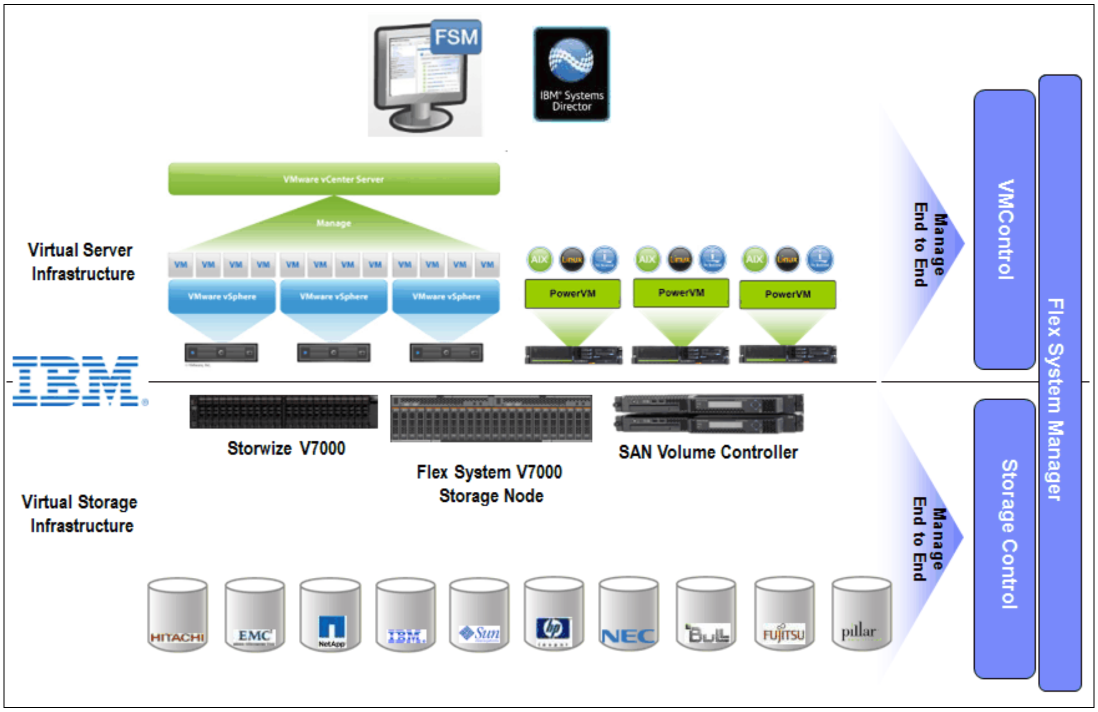


Figure 3-44 Managing diverse IT infrastructure with storage

For further details regarding managing storage using FSM, review the section 3.3.7, “Managing storage using IBM Flex System Manager” on page 139.

IBM Flex System Manager Editions

The IBM Flex System Manager software contains a set of components that are responsible for performing certain management functions. These components must be activated using the available IBM Feature on Demand (FoD) software entitlement licenses, and they are licensed on a per-chassis basis.

Note: The management node comes standard without any entitlement licenses, therefore, you must purchase a license to enable the required IBM Flex System Manager functionality.

IBM Flex System Manager base feature set

The following functionality and features are available at this time in the IBM Flex System Manager base feature set:

- Support for up to four managed chassis, and up to 5,000 managed elements
- Auto-discovery of managed elements
- Overall health status
- Monitor and availability
- Hardware management
- Security management
- Administration
- Network management (Network Control).Storage management (Storage Control)
- Virtual machine life cycle management (VMControl Express)
- I/O address management (IBM Fabric Manager)

IBM Flex System Manager advanced feature set

IBM Flex System Manager advanced feature set offers all capabilities of the base feature set plus:

- Image management (VMControl Standard)
- Pool management (VMControl Enterprise)

Important: The advanced upgrade license requires the IBM Flex System Manager base license.

IBM Fabric Manager (IFM)

IBM Fabric Manager (IFM) is available as a Feature on Demand (FoD) through the IBM Flex System Manager management software. It is an easy-to-use server provisioning I/O management tool. Unlike performing deployment steps manually, the parameters are configured automatically. IBM Fabric Manager monitors your data center and automatically responds to server issues, which reduces the risk of server failures and downtime.

IFM offers the following enhanced features and functions:

- Automatic provisioning of I/O addresses.
- Pre-defined configurations of I/O parameters, such as MAC addresses for Ethernet, World Wide Names and SAN boot targets - for up to 256 chassis or 3584 compute nodes.
- Provides I/O parameter and VLAN migration to standby compute nodes in case of compute node failure and executes action plans based on triggered events.
- Monitors the health of compute nodes, and automatically (without user intervention) replaces a failed compute node from a designated pool of spare compute nodes.
- Using an enhanced graphical user interface (GUI), creates addresses for compute nodes, saves the address profiles, and deploys the addresses to the slots in the same chassis (or in up to 256 different chassis).

IBM Fabric Manager enables you to quickly replace and recover compute nodes in your environment. This is done by assigning Ethernet MAC, Fibre Channel WWN, and SAS WWN addresses so that any compute nodes plugged into those bays take on the assigned addresses, which enables the Ethernet and Fibre Channel infrastructure to be configured once, and before any compute nodes are connected to the chassis.

After receiving a failure alert, IFM attempts to power off the failing compute node, reads the IFM virtualized addresses and boot target parameters, applies these parameters to the next compute node in the standby pool, and power on the standby compute node.

Additionally, you can create profiles for chassis that have not been installed in the environment by simply associating an IP address to the future chassis.

You must configure the CMM to enable IFM access. The IFM configuration is chassis-based and does not transfer with the physical CMM. When a CMM is moved to a new chassis, it clears out its IFM configuration, and the IFM configuration must be reapplied on the new CMM.

Important: You must set the TCP Command Mode Protocol on the CMM Web interface.

There are minimum hardware and software requirements the system must meet before you can install or use IFM, which are documented in detail at:

http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.iofm.doc/dw1li_creating_a_configuration_file_automatically.html

For information regarding how to install, configure, and use IBM Fabric Manager (IFM), visit http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.iofm.doc/dw1li_advanced.html

3.3.3 Logging in to IBM Flex System Manager Node

The following section gives step by step instructions regarding how to log in to the FSM.

1. Open a web browser, and type in the IP address of the FSM. Remember that depending on the security style, it might be needed to use https protocol
2. At the Login screen, enter your User name and Password, as shown in Figure 3-45

Figure 3-45 Shows the FSM login page.

Figure 3-45 IBM Flex System Manager login screen

3. After a successful login, it takes you to the default Home page of IBM Flex System Manager. Figure 3-46 on page 117 shows the FSM Home page.

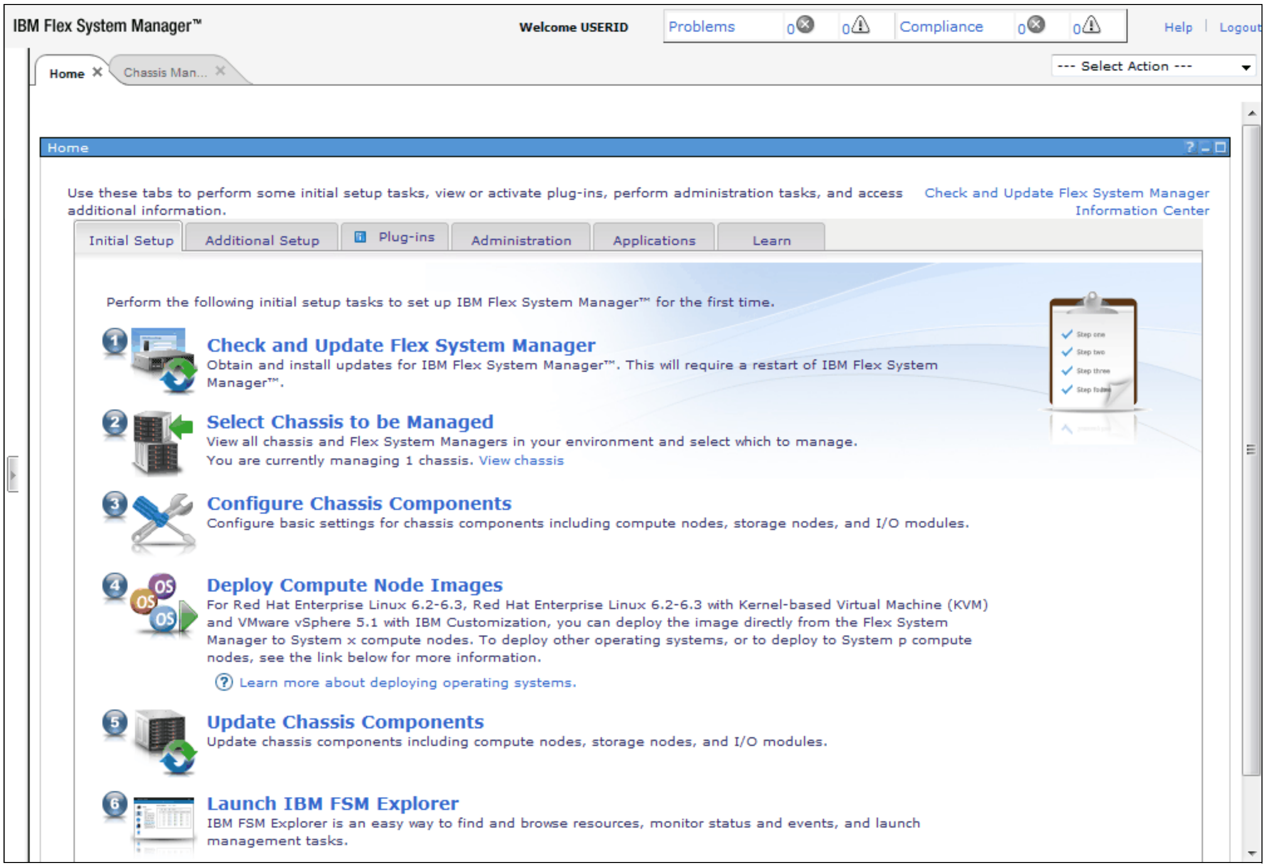


Figure 3-46 FSM - Startup Page

3.3.4 Overview of IBM Flex System Manager and IBM FSM Explorer

The new GUI on IBM Flex System Manager provides 2 interfaces, which will be discussed in the following sections.

IBM Flex System Manager

Following section will give a short overview on how to navigate around in the IBM Flex System Manager.

Home page as shown above in Figure 3-46, has six tabs which helps you with setting up and administer the system.

Home tab Initial setup of the system

Additional setup tab Helps to complete initial setup and prepare IBM Flex System Manager™ for production

Plug-ins tab Different functions added to the FSM, such as Storage and Network control

Administration tab A collection of links to ease up configuration and administration

Applications tab Additional applications added to FSM for more functions, such as IBM Fabric Manager

Learn tab Links to online education videos regarding the usage of IBM Flex System

Chassis manager page has multiple functions as shown in Figure 3-47.

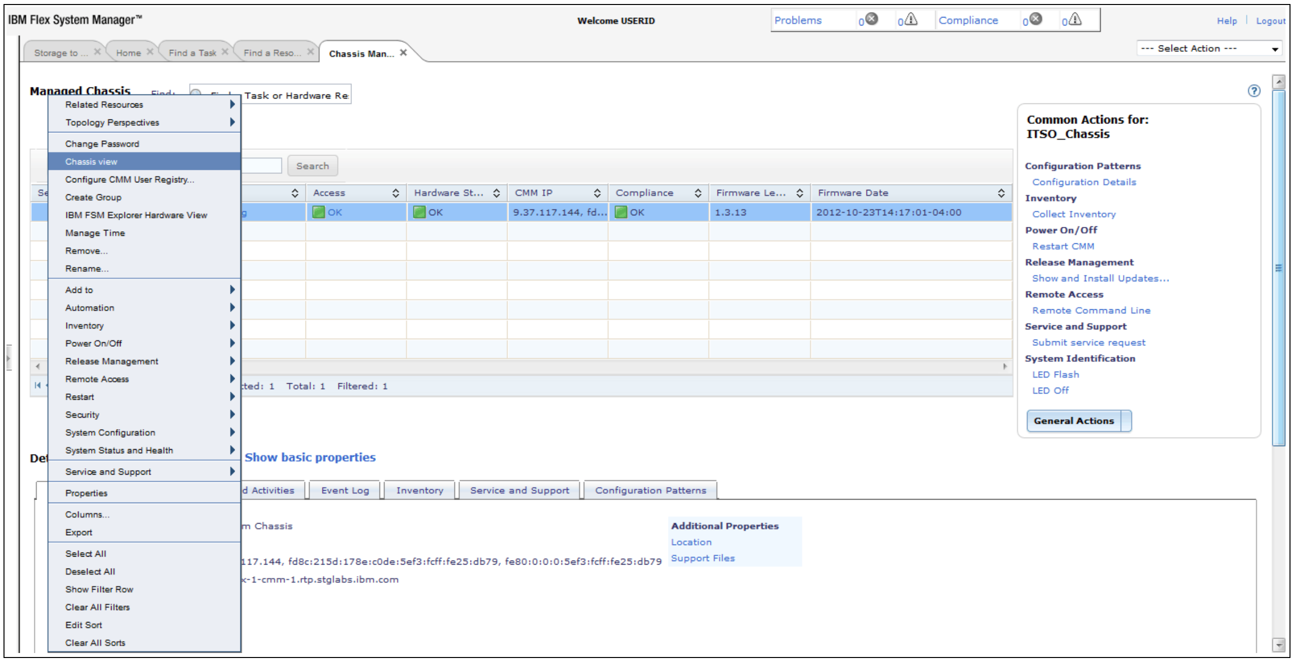


Figure 3-47 FSM chassis manager tab

Figure 3-47 shows FSM chassis manager. In this window,

- Clicking on the chassis name, will open IBM FSM Explorer, which will be discussed later in this section.
- The global search tab, which helps finding tasks and hardware resources
- The action tab, gives you advanced configuration options of the chassis
- The details section gives you detailed information regarding the chassis, such as eventlog and inventory.

In the FSM GUI, there is a small slide-out tab in the center-left of the interface. Clicking on it opens a menu, which shows a comprehensive list of functions that FSM offers. Figure 3-47 shows this FSM slide-out menu.

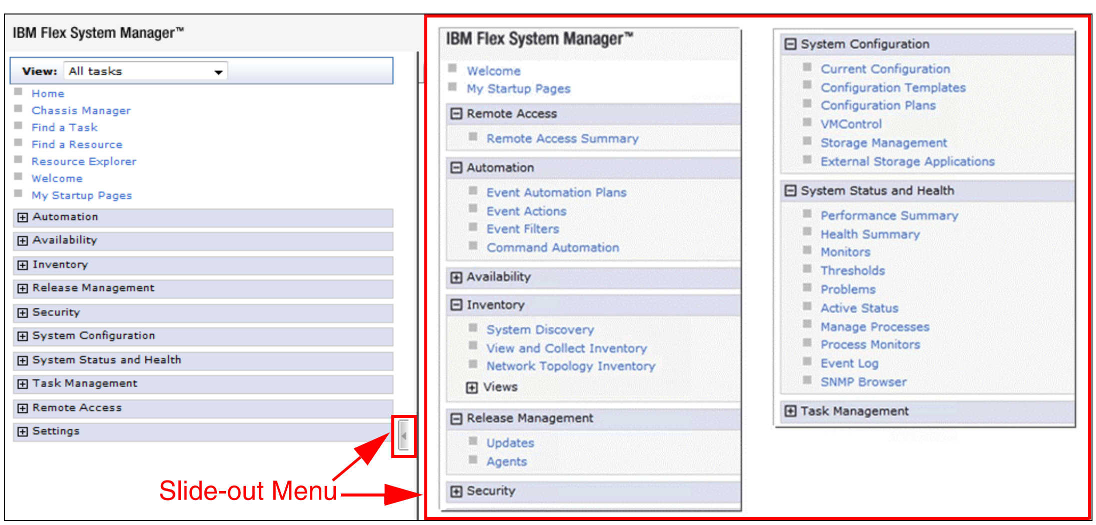


Figure 3-48 FSM - Slide-out menu

As seen in Figure 3-48 on page 118, there are numerous enhanced features and functions that can be easily preformed in the user-friendly FSM user-interface. For detailed information regarding IBM Flex System Manager's web interface and management options, refer to the IBM Flex System Manager Systems Management Guide, at
http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.director.main.helps.doc/fsm_1_fqr0_bk_systm_mgmt.pdf

IBM FSM Explorer

Following section will give a short overview on how to navigate around in the IBM FSM Explorer

IBM FSM Explorer has a horizontal pull down menu, and a vertical menu. The horizontal menu, is divided into two menus, one in the top black frame, showing health and job status, and the second in the blue frame, which guides you to different tasks and hardware recourses.

To get to IBM FSM Explorer, click on "Launch IBM FSM Explorer" in the home tab of IBM Flex System Manager, this will open IBM FSM Explorer in a new tab. Figure 3-49 shows IBM FSM Explorer home page.

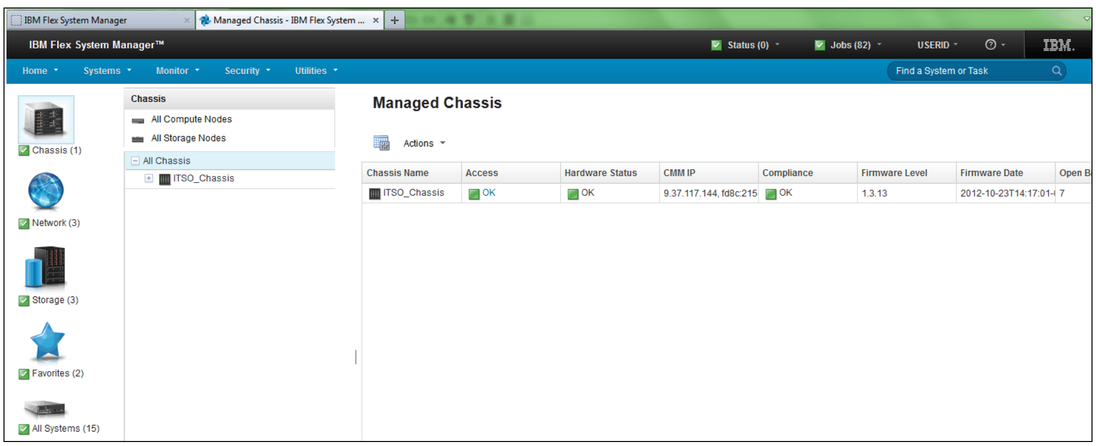


Figure 3-49 IBM FSM Explorer home - tabs

When working in IBM FSM Explorer, it is important to know that some tasks will be opened in IBM FSM Explorer, and other tasks will be opened in IBM Flex System Manager. When choosing tasks in the horizontal menu, it will inform you if the selected task is opened in IBM Flex System manager. Figure 3-50 shows how to identify where a chosen task will open.

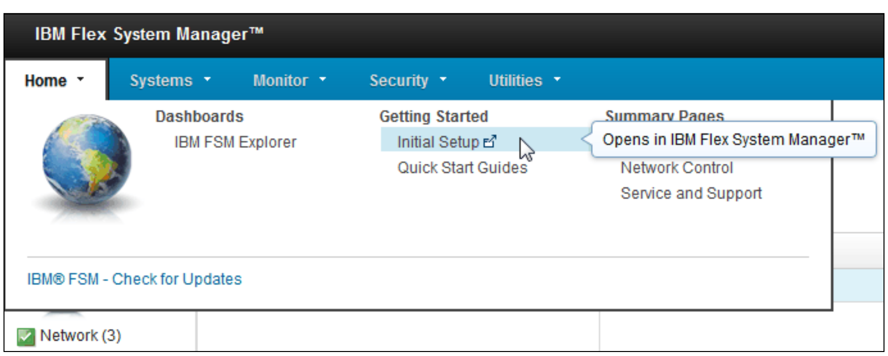


Figure 3-50 IBM FSM Explorer - Tasks

IBM FSM Explorer - Important information: If you choose a task from the chassis hardware map, or any other menu than the horizontal menu, you will not be informed if it is opened in IBM FSM Explorer or in IBM Flex System Manager.

Horizontal menu

The horizontal menu is divided in 5 pull down menus, which we will cover in this section.

Home menu is divided in to 3 groups:

- Dashboard takes you back to the main page.
- Getting started gives you shortcuts to initial setup off the FSM and quick start guides in Info Center.
- Summary pages gives you shortcuts to Storage control, Network control and service & support pages in IBM Flex System Manager.

Figure 3-51 shows the Home menu.

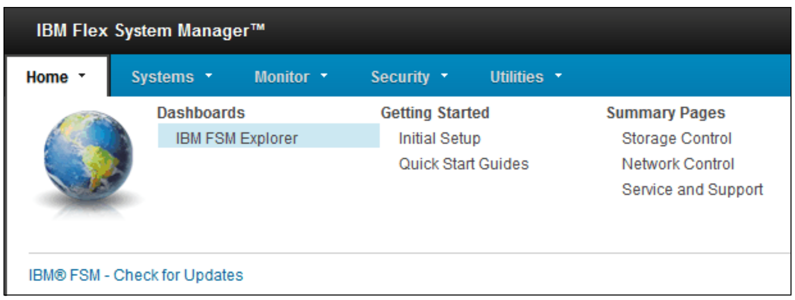


Figure 3-51 IBM FSM Explorer - Home menu

Systems menu is divided in to 3 groups:

- Systems gives you shortcuts to all groups defined, the OS group and Power systems management. Power systems management shows all power systems recourses.
- Configuration takes you to configuration patterns.
- Firmware takes you to update manager in IBM Flex System Manager.

Figure 3-52 shows the Systems menu.

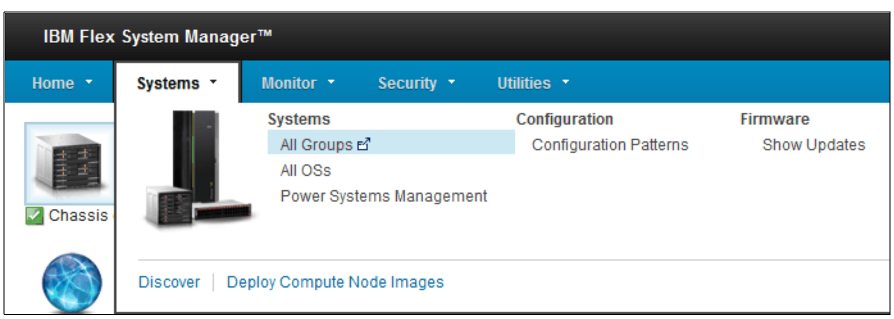


Figure 3-52 IBM FSM Explorer - Systems menu

Monitor menu is divided in to 5 groups:

- Status gives you shortcut to Active Status page, showing open alerts, Event log page and Service problems page.

- Monitoring gives you shortcuts to Monitors page and Threshold page, which allows you to create thresholds, and select which systems these thresholds should be used for monitoring.
- Compliance gives you shortcuts to Update manager (Monitored systems) and Compliance policy page.
- Automation gives you shortcuts to the event automation plans page, Event actions page, Event filters page and Command automation page.
- Jobs gives you a shortcut to the active and scheduled jobs page.

Figure 3-53 shows the Monitor menu.

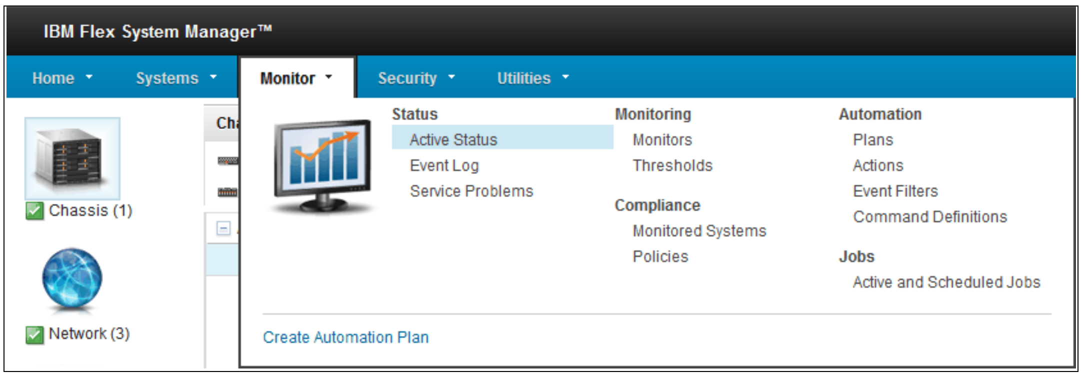


Figure 3-53 IBM FSM Explorer - Monitor menu

Security menu is divided in to 2 groups:

- Permissions gives you shortcuts to Users and Groups, Roles and Credentials which is used to give access and rights.
- Access takes you to Systems with no access, which is recourses, which there have not been granted access to.

Figure 3-54 Shows the Security menu.

Figure 3-54 IBM FSM Explorer - Security menu

Utilities menu is divided in to 3 groups:

- Settings gives you links to Auto-rename page which allows you to rename recourses, Service and support page, which allows you to configure the system to automatically report errors to IBM and Server Auditing, Which takes you to credentials page.
- Administration gives you links to Administration, Plug-ins and Application tab in IBM Flex System Manager Home page.
- User Preferences gives you links to FSM server and network management.

Figure 3-55 shows the Utilities menu.

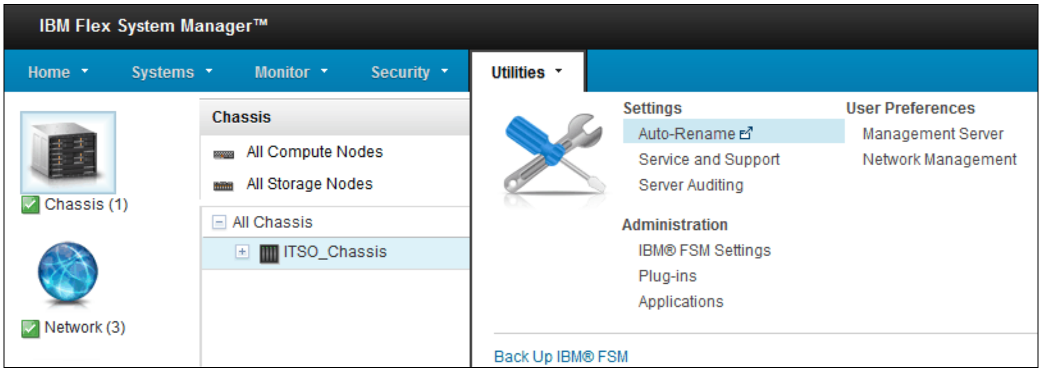


Figure 3-55 IBM FSM Explorer - Utilities menu

Vertical menu

The vertical menu has 5 images, which opens different pages, these pages will be covered in the following section.

Figure 3-56 Shows the vertical menu.

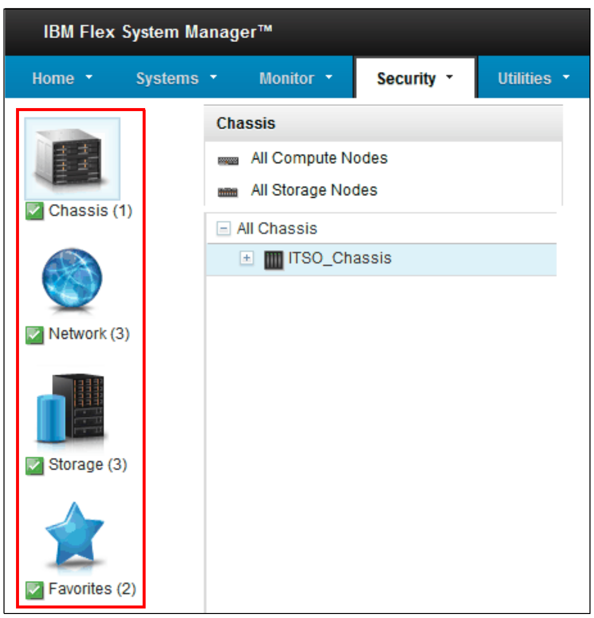


Figure 3-56 IBM FSM Explorer - Vertical menu

Chassis page gives you an overview of what is installed in one or more chassis as shown in Figure 3-57 on page 123.

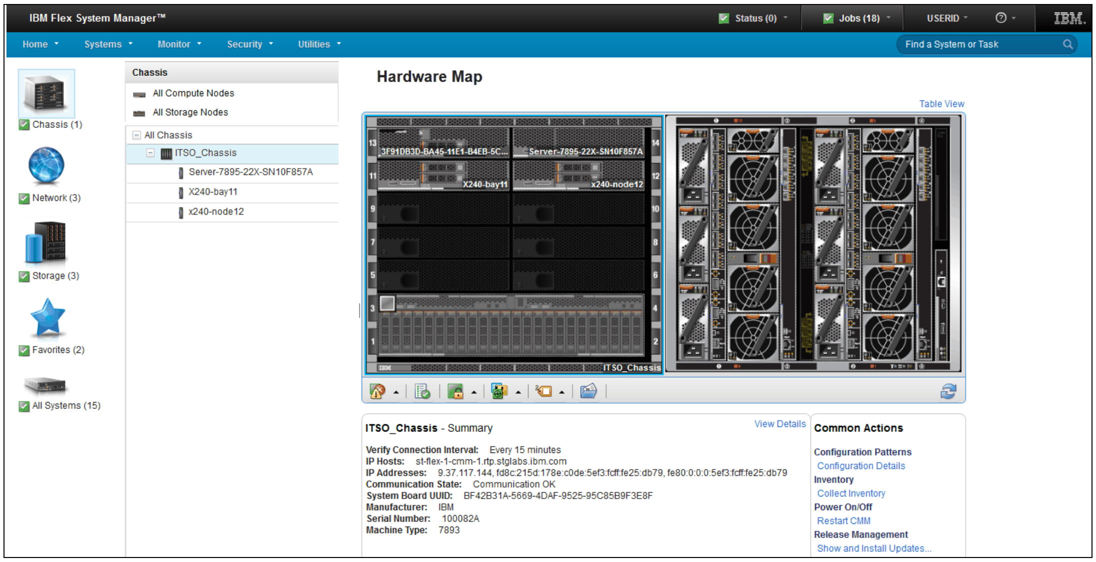


Figure 3-57 IBM FSM Explorer

To get to the chassis map, click on chassis in the vertical menu on Figure 3-56 on page 122, then click on the chassis under the chassis list as shown in Figure 3-58.

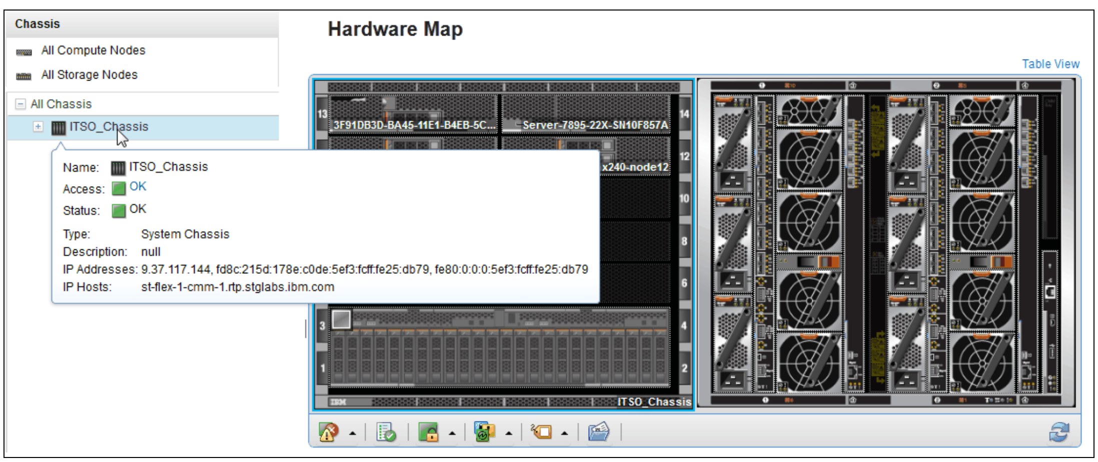


Figure 3-58 IBM FSM Explorer home tab - chassis map

The chassis view shows a map of the front and back with graphical representation of the components, as shown in Figure 3-59 on page 124.

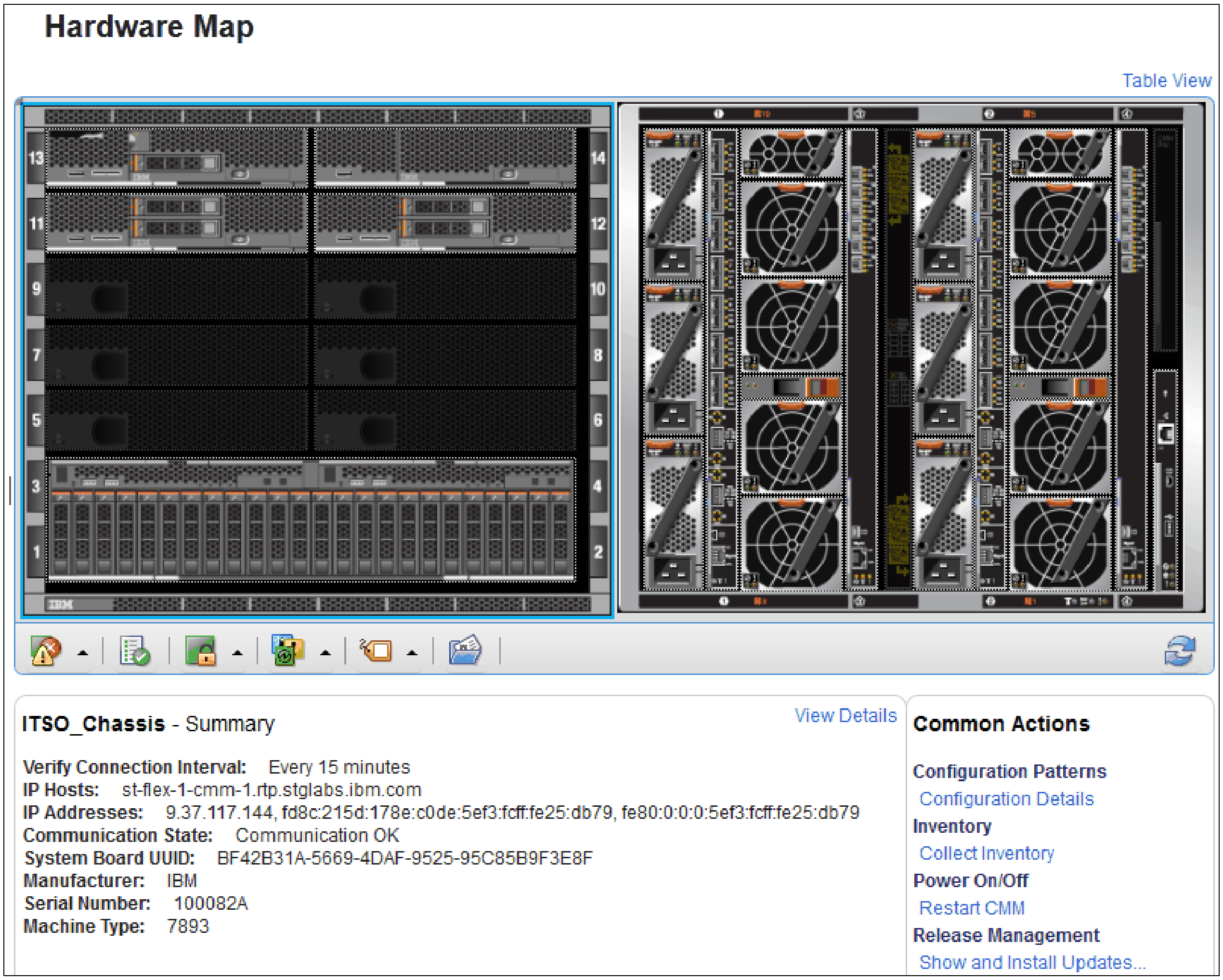


Figure 3-59 IBM FSM Explorer chassis map

To get detailed information about the chassis and chassis components, click on **View Details** in the summery field of Figure 3-59.

As shown in Figure 3-60, *Common Actions* for chassis are shown below the Hardware map at the left side and the detailed information regarding the chassis are shown below the Hardware map at the right.

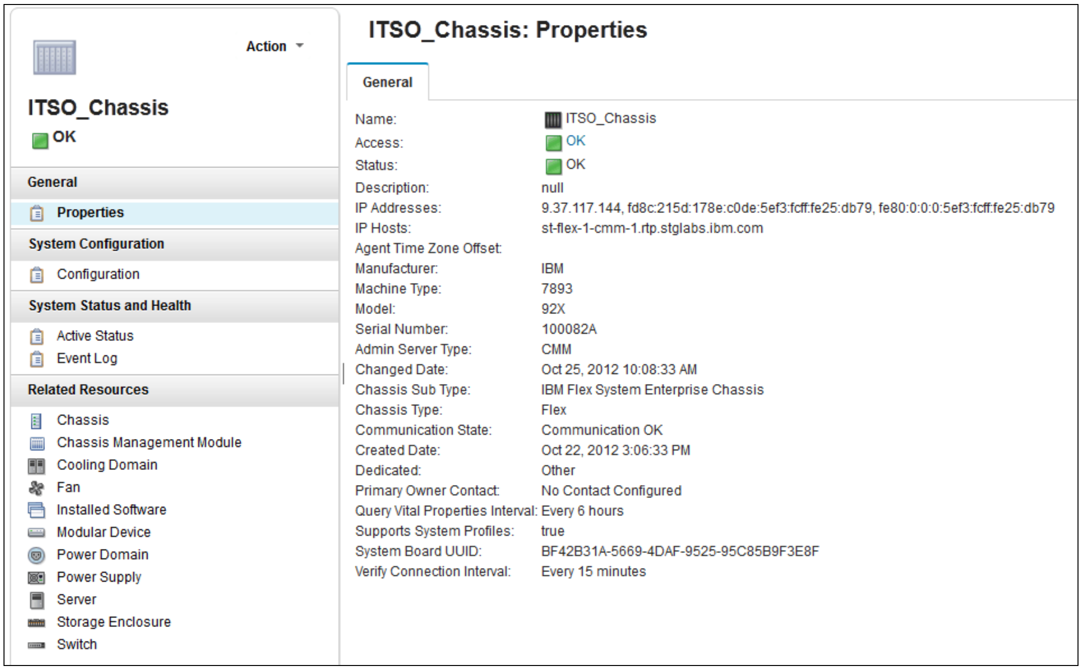


Figure 3-60 IBM FSM Explorer chassis details

A drop-down menu shows the actions that can be performed at the Chassis level, as shown in Figure 3-61.

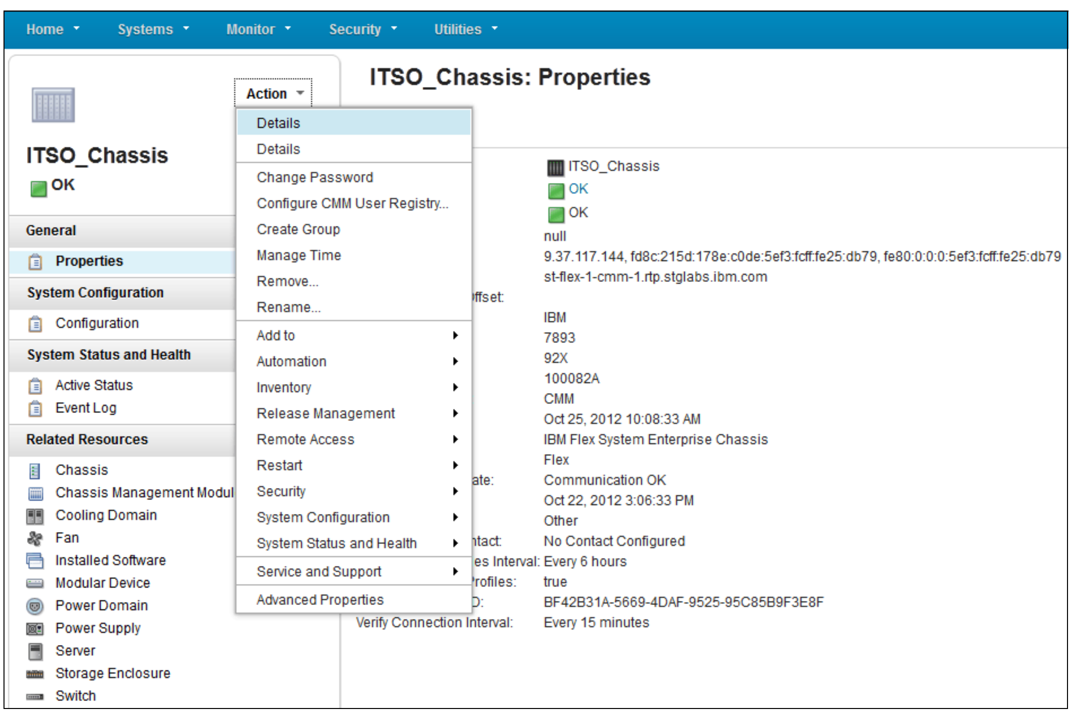


Figure 3-61 IBM FSM Explorer - Component Actions

Network page gives an overview of VLANs, Subnets and network switches, as shown in Figure 3-62.

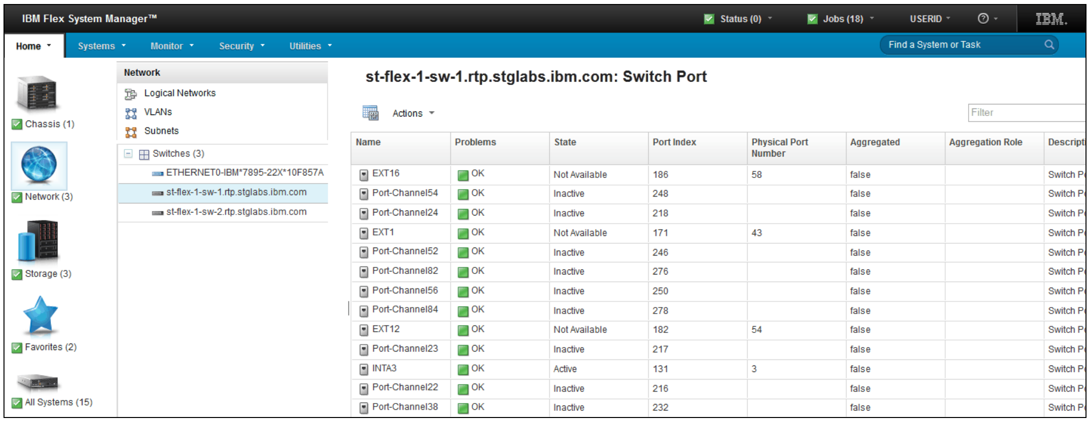


Figure 3-62 IBM FSM Explorer - Network page

Storage page gives you an overview of all storage systems, Fabrics and Storage pools as shown in Figure 3-63.

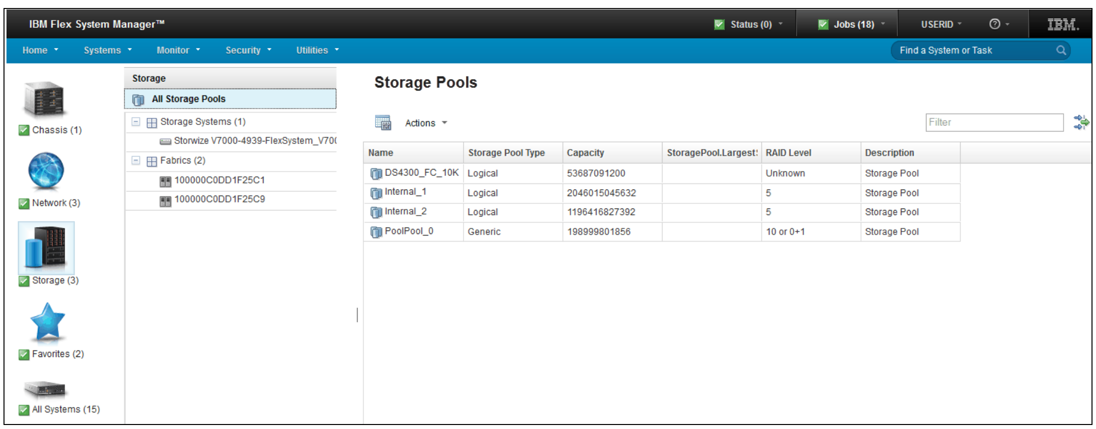


Figure 3-63 IBM FSM Explorer - Storage page

Right clicking on an IBM Flex System V7000 Storage Node, gives you the possibility to update firmware as shown in Figure 3-64 on page 127.

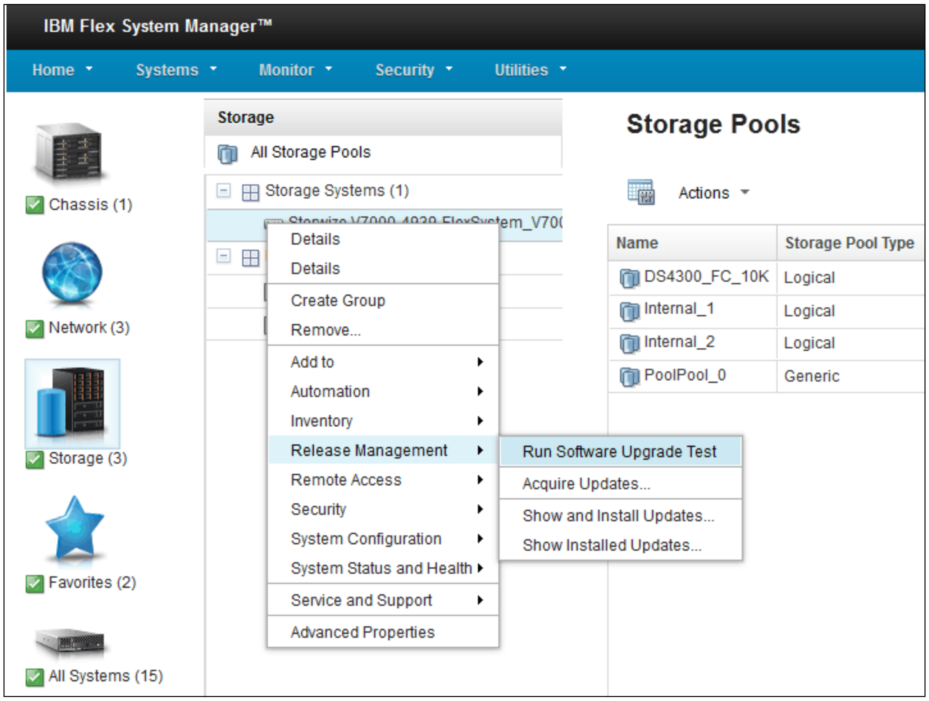


Figure 3-64 IBM FSM Explorer - Storage page - Firmware update

From the menu shown in Figure 3-64, you can also launch the IBM Storwize V7000 Element manager as seen in Figure 3-65.

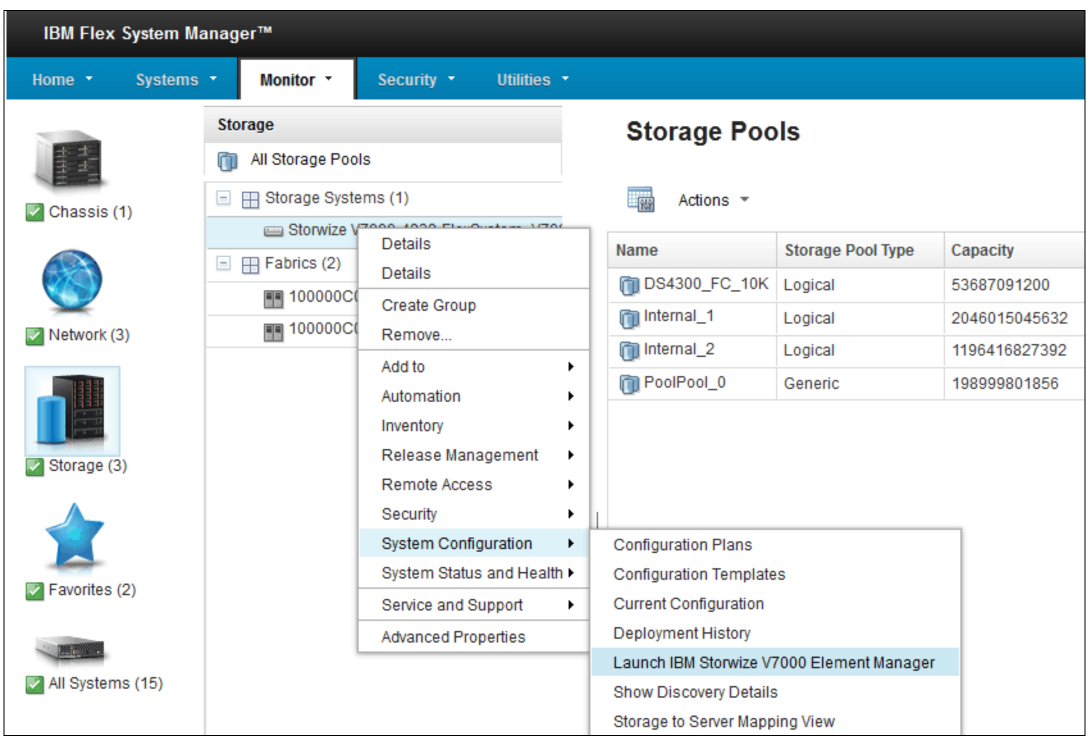


Figure 3-65 IBM FSM Explorer - Storage page - launch IBM Storwize V7000 Element manager

Choosing the option to Launch the IBM Storwize V7000 Element Manager will open the IBM Flex System V7000 Storage Management login page, as shown in Figure 3-66 on page 128.

Figure 3-66 IBM Flex System Storage Management - Login page

In the Fabrics section, clicking on one of the numbers (Last 12 numbers is the MAC address of the switch) will show you all the switches as shown in Figure 3-67.

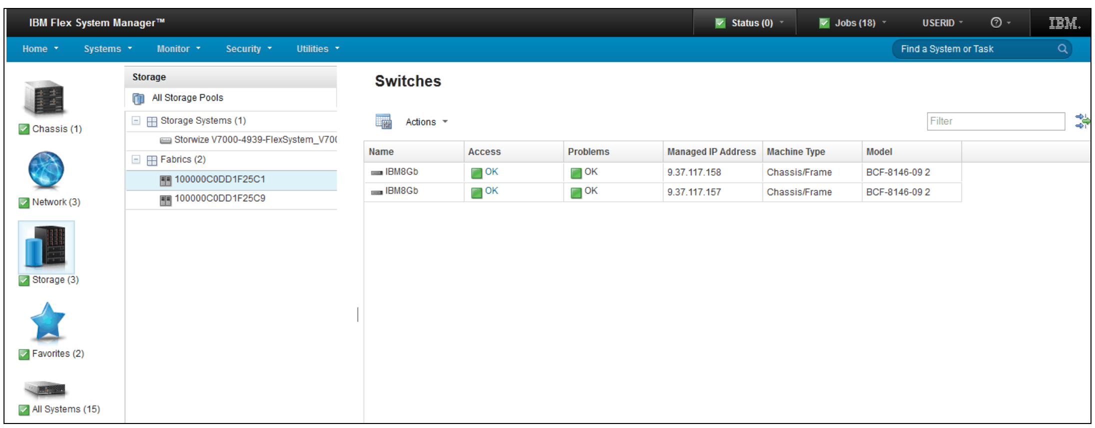


Figure 3-67 IBM FSM Explorer - Storage page - Switch overview

Double clicking on one of the switches leads you to detailed information about the switch, as shown in Figure 3-68 on page 129.

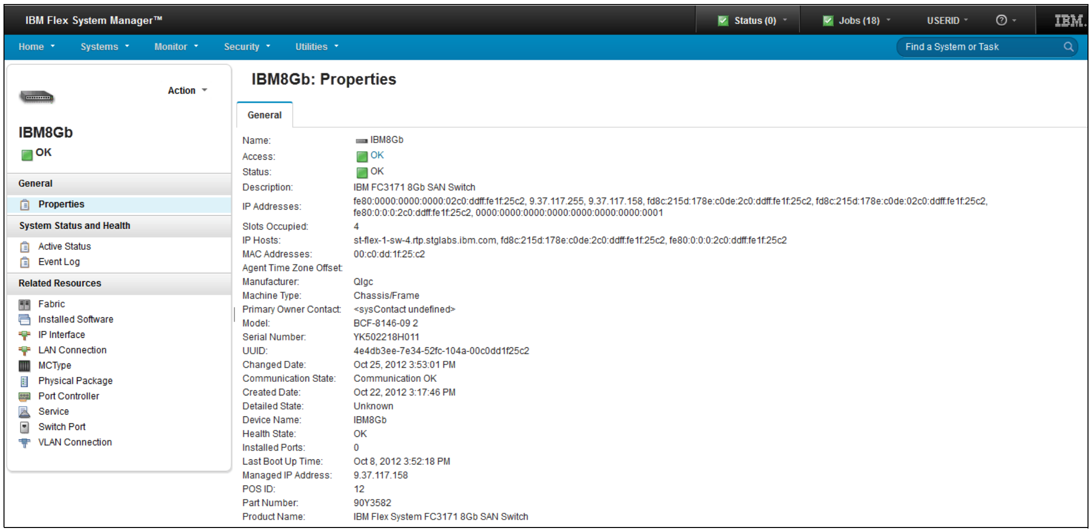


Figure 3-68 IBM FSM Explorer - Storage page - Switch details

From the switch overview page, right click on one of the switches select **Remote access** → **Launch Web Browser**. This will open the switch web GUI login prompt, from where the switch can be configured. Figure shows how to open the switch web GUI.

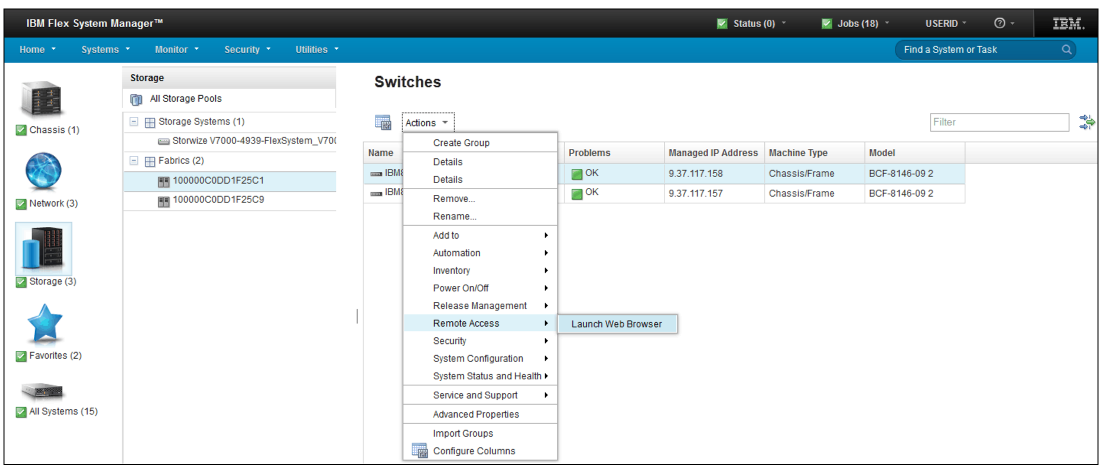


Figure 3-69 IBM FSM Explorer - Storage page - Switch GUI

Figure 3-70 Shows the FC switch login page.

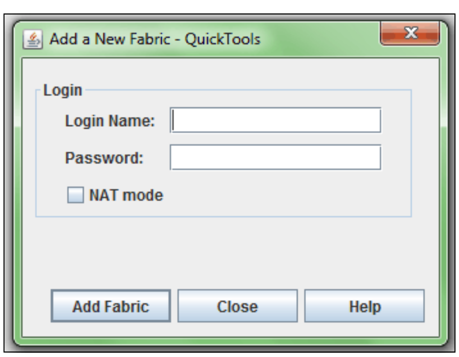


Figure 3-70 FC Switch login

3.3.5 Accessing I/O Modules using FSM

In an IBM Flex system, node to node communication happens within the Chassis.

Network Control

Network Control is a plug-in that enables utilization of integrated network management capabilities of an FSM. It has the following key features:

- Logical network management – allows management of port profiles, VLANs, ACLs and QoS in virtualized, live-migration environments
- Leverages 802.1Qbg standards in integrated switches and PowerVM, KVM and IBM "Osiris" vSwitch for VMware
- Optional Fabric Management extends QoS Management providing advanced monitoring, VM priority and rate limiting
- Network monitoring at a glance via network topology perspectives with the ability to see the components affected by network outages
- Enable end-to-end network and virtualization management

To access Network Control, from IBM FSM Explorer, select **Home menu** → **Network Control**, as shown in Figure 3-71.

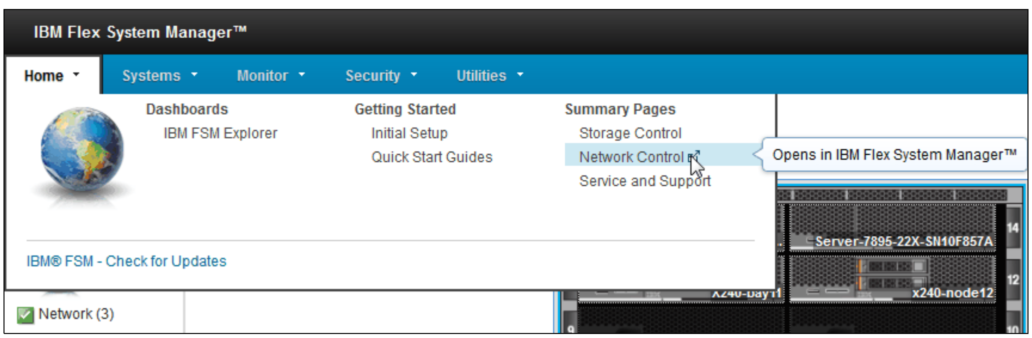


Figure 3-71 IBM FSM Explorer - Network control

This will open up the Network Control interface, which shows management and configuration options for the IBM Flex System I/O Modules, as shown in Figure 3-72 on page 131.

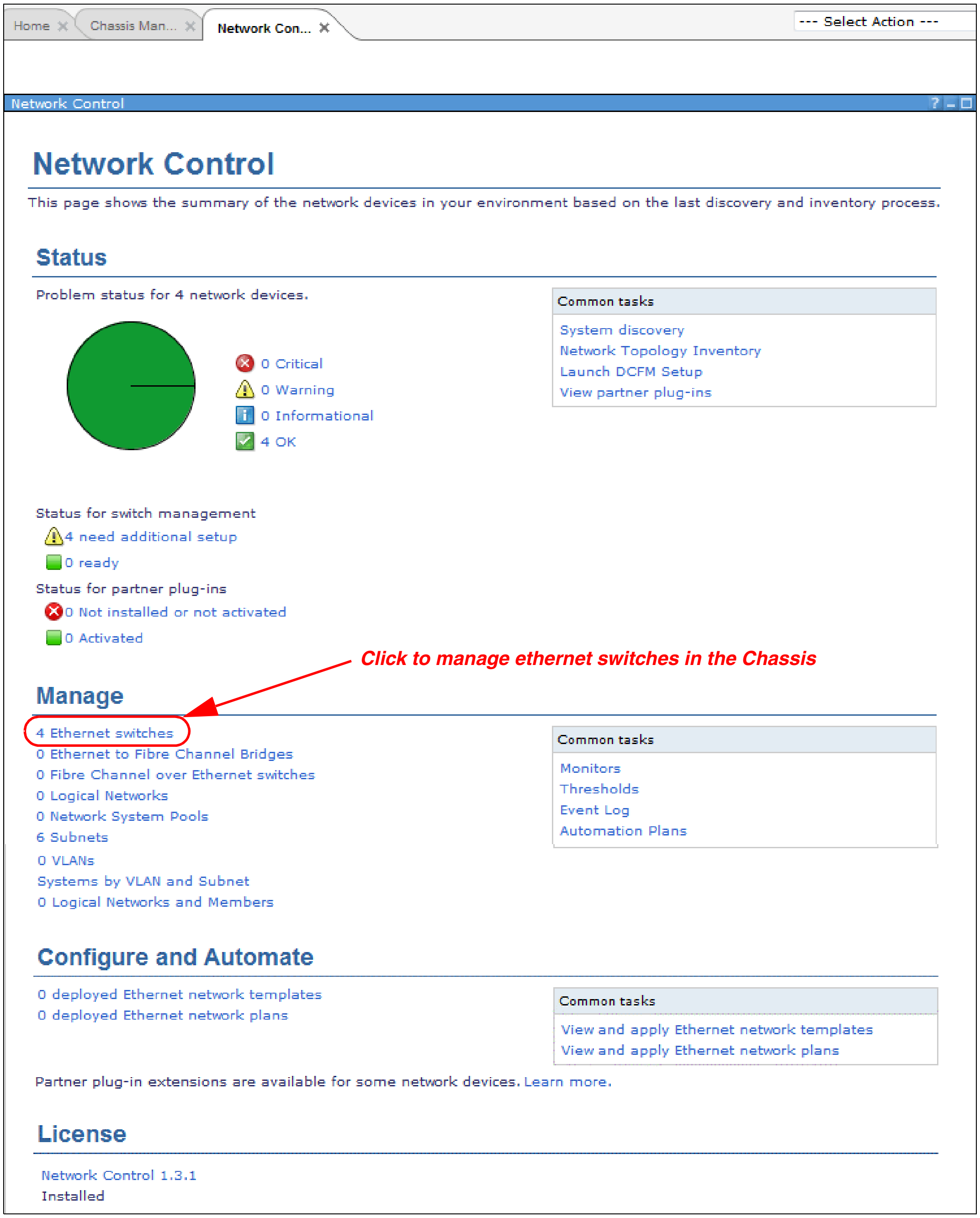


Figure 3-72 Network Control Plug-in

Selecting **Ethernet Switches** in Network Control interface opens a view of all available Ethernet Switches along with information regarding state, access, IP and type, as shown in Figure 3-73 on page 132.

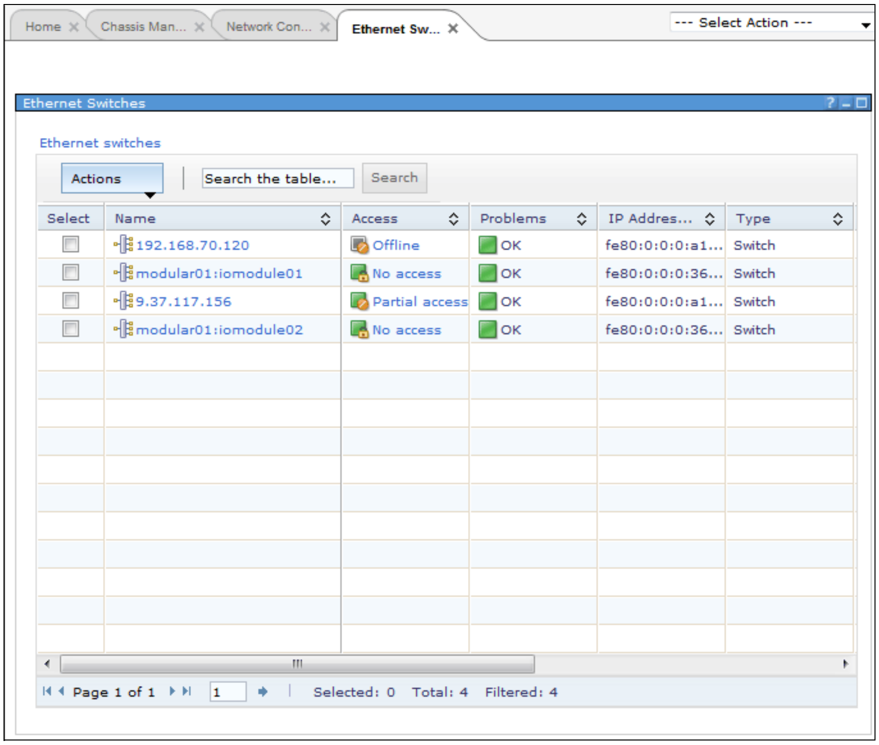


Figure 3-73 Viewing Ethernet Switches in FSM

Select a switch by clicking on its given *name* to go to the detailed management and configuration options for it, as shown in Figure 3-74 on page 133.

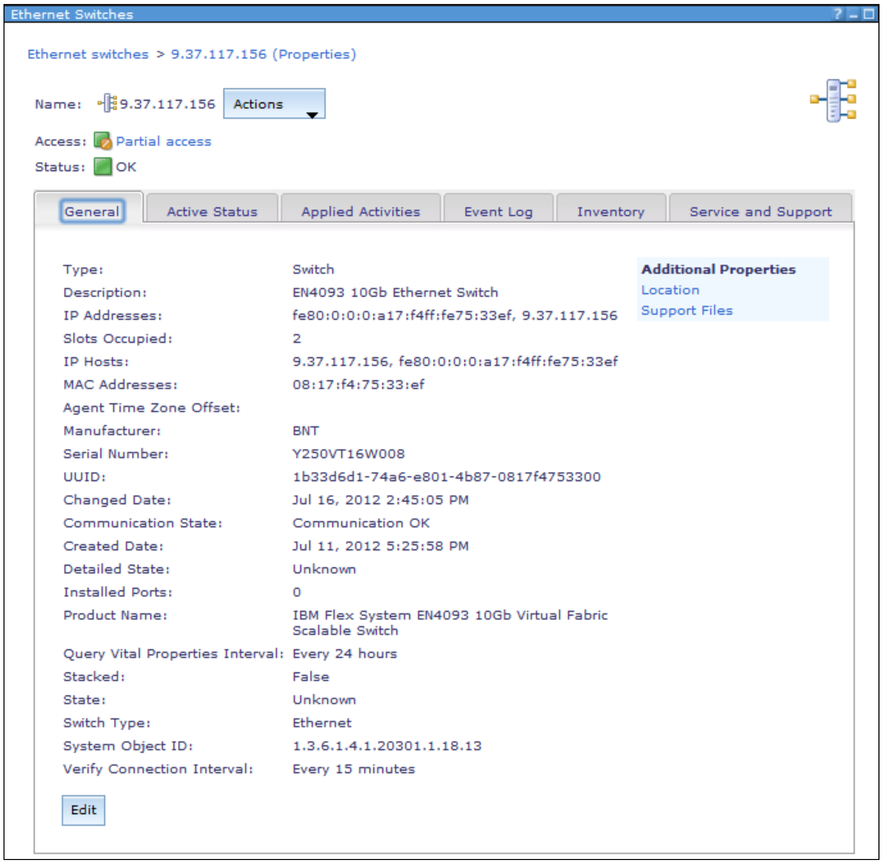


Figure 3-74 Details regarding selected ethernet switch

Several tabs in this view illustrate enhanced monitoring and logging capabilities of the FSM. Click on the *Inventory* tab to view and gather Inventory for the selected switch module, as shown in Figure 3-75 on page 134.

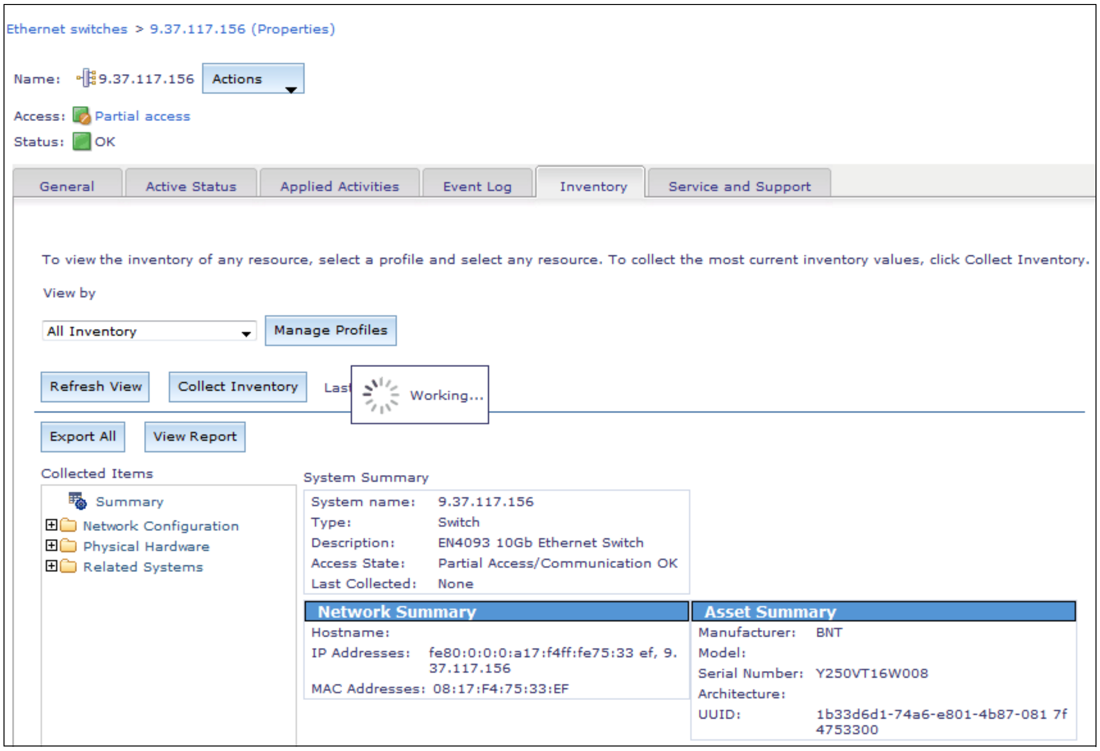


Figure 3-75 Gathering Inventory information

For further details regarding the management and configuration of I/O Modules using IBM Flex System Manager, visit
http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.8731.doc/managing_network_resources.html

3.3.6 Data Collection using FSM

You can use IBM Flex System Manager to collect and submit support files for a managed resource. Support files can contain detailed system information used to help diagnose a serviceable hardware problem, dump files collected from a managed resource, event logs, and more.

By default, a plug-in FSM tool called Service and Support Manager automatically collects additional data associated with a serviceable hardware problem, and stores it as a support file. However, you can also collect support file data manually, even when a serviceable hardware problem has not occurred.

Service and Support Manager (SSM)

IBM Service and Support Manager monitors, tracks, and captures system hardware errors and service information, and reports serviceable problems directly to IBM Support using the IBM Electronic Service Agent (ESA) tool. The IBM Flex System Manager will automatically monitor for any serviceable problems on your managed hardware and automatically report problems to IBM service.

IBM Service and Support Manager offers enhanced reporting abilities. It opens service requests for failed hardware using embedded Electronic Service Agent (ESA). It also collects, transmits, and manages extended error data including logs and diagnostics. Another added

advantage of SSM is transmission of hardware and software inventory data to IBM using ESA, which assists in troubleshooting the problem as well as monitor the dependencies.

Note: If you do not setup the Electronic Service Agent (ESA) then problems will be collected in the FSM management appliance but will not be reported to IBM, however, you can always use the FSM itself to monitor and track your serviceable problems.

To access the Service and Support Manager (SSM), click on **Home menu** → **Service and Support**, as shown in Figure 3-76.

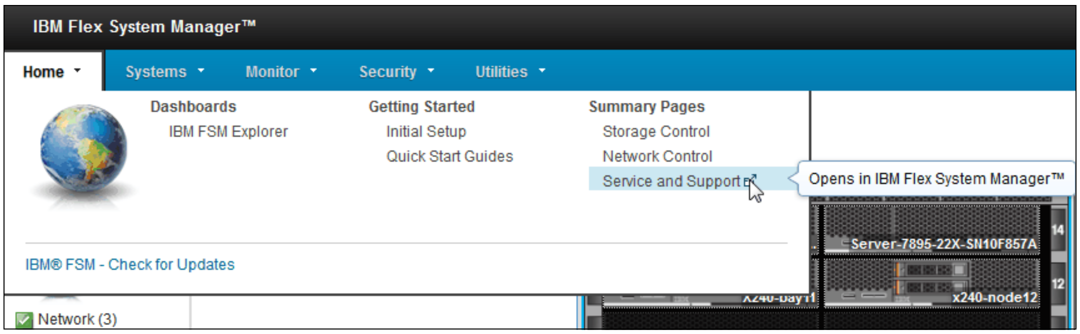


Figure 3-76 Opening Service and Support Manager (SSM)

SSM interface opens up, which shows comprehensive set of Service Management tools in one tabular view, as shown in Figure 3-77 on page 136.

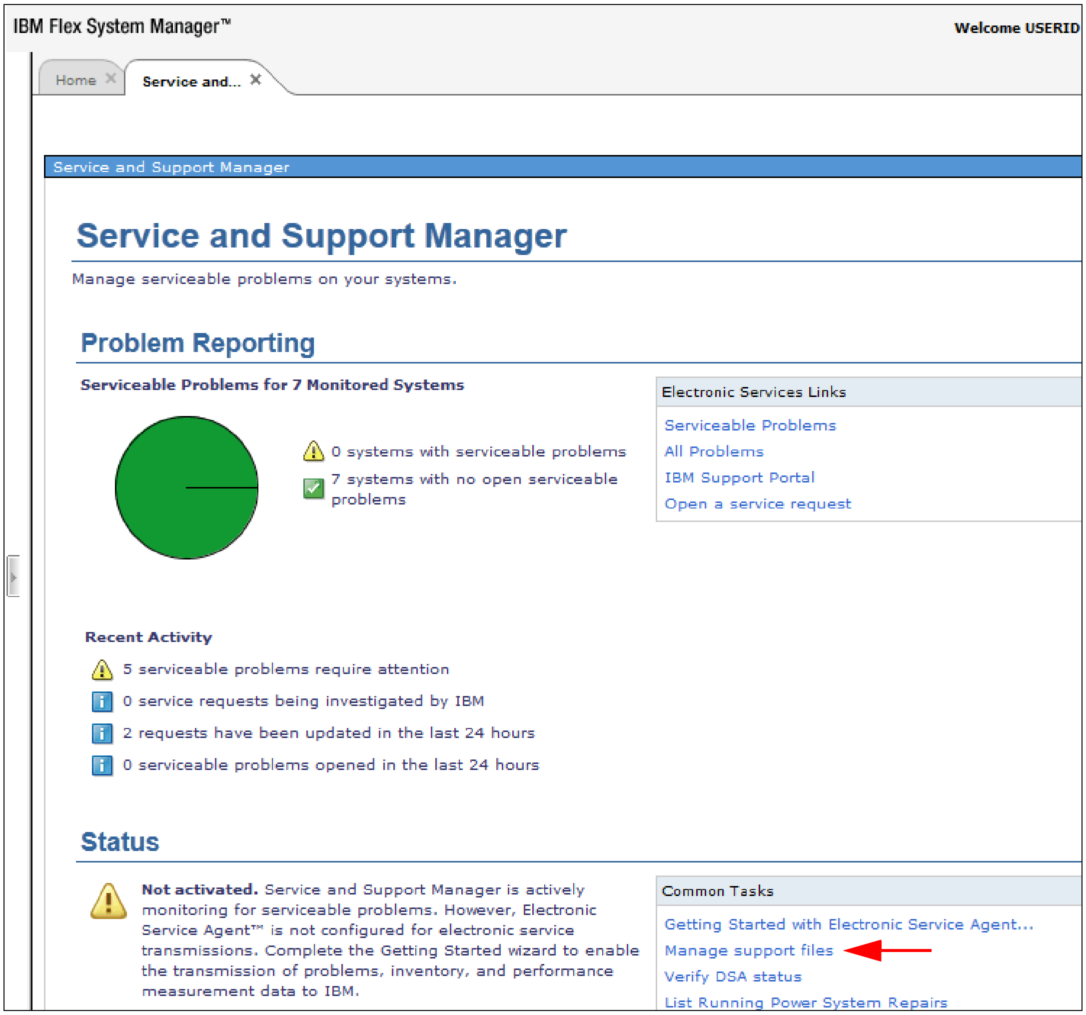


Figure 3-77 SSM default view

To collect logs from IBM Flex System V7000 Storage Node using IBM Flex System Manager, click on **Manage support files** as shown in Figure 3-77.

Manage support files page comes up and by choosing **Collect support files**, a new window appears, where you can see logs already collected. Figure 3-78 on page 137 shows logs already collected.

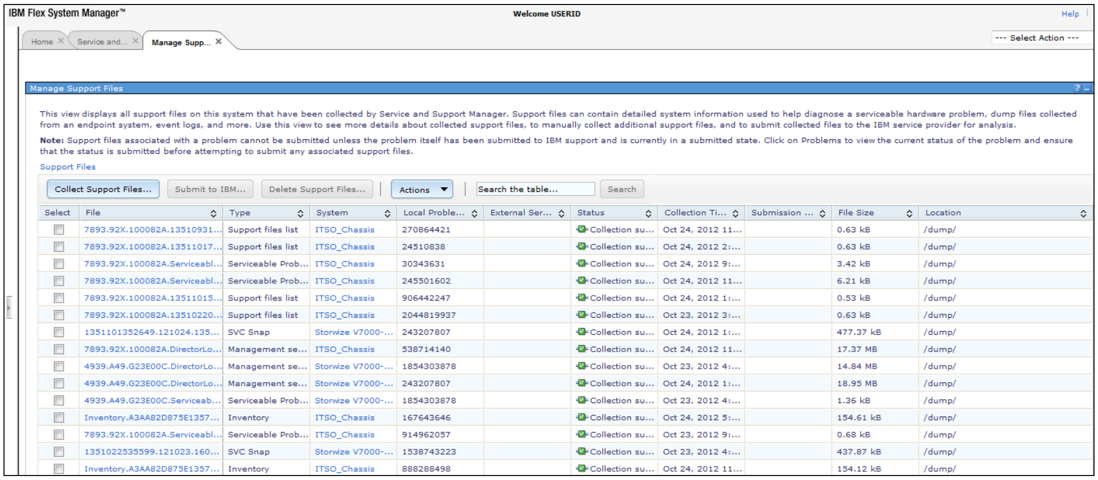


Figure 3-78 IBM FSM - Manage support files

By selecting one or more files, then click actions, and you have several options to download or transfer the logs:

- Submit to IBM (only works if Electronic Service Agent is enabled)
- Upload to ftp server: Upload to a ftp server on your network
- Copy to media: USB key, inserted into IBM Flex System Manager Node
- Download: Download the support file to the computer which are connected to the IBM Flex System Manager

Figure 3-79 Shows the options for a support file.

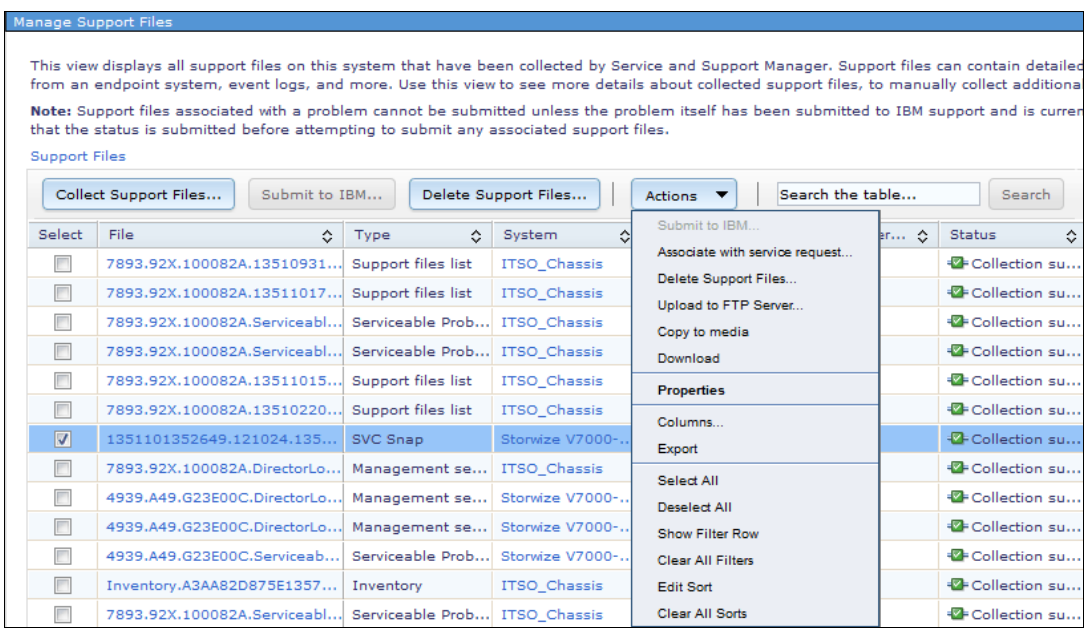


Figure 3-79 IBM FSM - Manage support files download

To collect a new support file, click **Collect Support Files**. A new window will appear, where you choose the device as shown in Figure 3-80 on page 138.

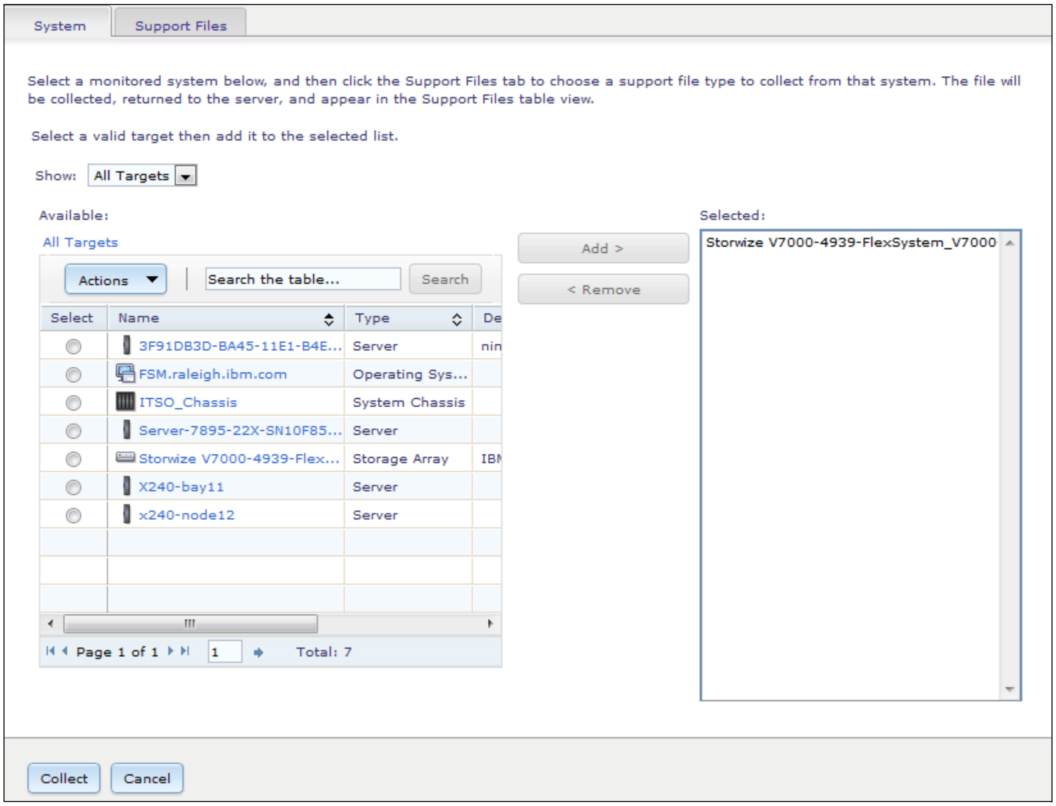


Figure 3-80 IBM FSM - Collect support file - choose device

To select the type of support file you which to collect, click the *Support Files* tab, then select the file and click *collect* as shown in Figure 3-81.

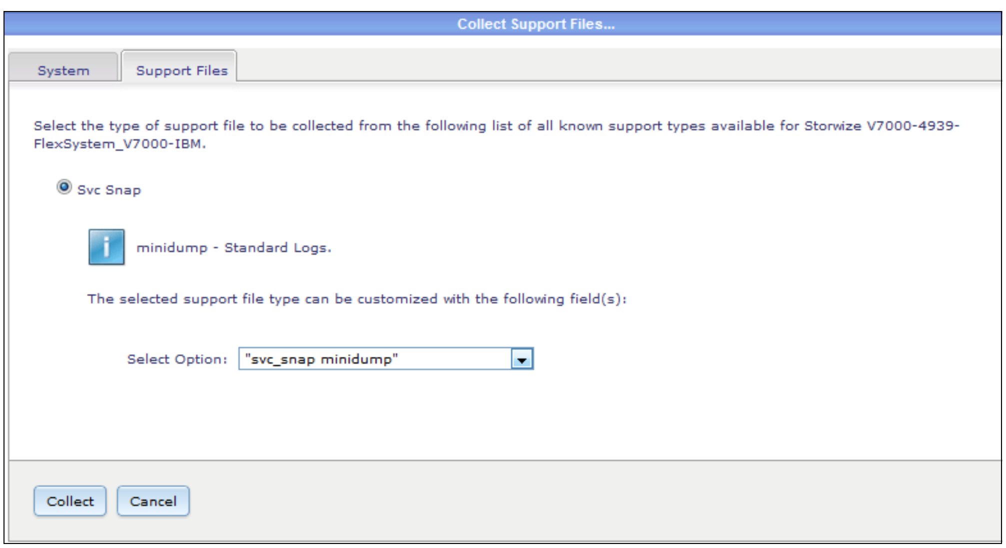


Figure 3-81 IBM FSM - Collect support file - choose file type

When the file has been collected, you can choose to download it, send it to IBM or copy it to a media.

3.3.7 Managing storage using IBM Flex System Manager

Storage management with IBM Flex System Manager involves two software components: Storage Manager and Storage Control. Both components are included with the management software.

- Storage Manager is a standard management software capability that provides basic storage life cycle management (Discovery, Inventory, Health and Alerts).
- Storage Control is a pre-installed plug-in for the management software that expands storage support to mid-range and high-end storage devices.

The Storage Control feature of IBM Flex System Manager allows management of most IBM mid-range and high-end storage systems. With the Storage Control plug-in, you can manage an expanded set of storage subsystems and Fibre Channel switches.

Storage Control offers the following key advantages:

- Extends storage management of IBM Systems Director and VMControl to cover most IBM storage systems
- Storage device discovery and coverage in integrated physical and logical topology views
- Shows relationships between storage and server resources
- Facilitates single management server deployment
- Provides ability to view controller and volume status and to set notification alerts
- Virtualize for higher storage utilization
- Balances high performance and cost for mixed workloads
- Protects data and minimizing downtime
- Integration with VMControl featuring the following additional functions:
  - Storage provisioning for image creation, deployment, and cloning
  - Ability to manage storage system pool life cycle, take group actions across pool and policy-based storage placement, provisioning, and cloning actions within the pool

Storage Control also extends the management capabilities and device coverage of the VMControl, which is a tool within the IBM Flex System Manager Node responsible for the life cycle management of virtual resources. This collaboration is leveraged in VMControl for tasks like working with system pools for items like server and storage pool creation, the image library management and storage, and integrated virtualization management across server and storage infrastructure.

For further details regarding storage management features, visit
http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.director.storage.helps.doc/fqm0_t_sm_configuring_storage.html

Storage Management and Storage Control

You can use Storage Control to monitor the health of storage subsystems, storage pools, storage volumes, and fabric switches. You can run inventory against the Storage Control farm to see updated health status.

You can also use Storage Control to manage several external storage subsystems. These subsystems include the IBM® DS3000, DS4000®, DS5000, families, the IBM DS8000® family (DS8100, DS8300, DS8700, DS8800), the IBM SAN Volume Controller, the IBM Storwize V7000, IBM XIV®, Qlogic 8GB Fibre Channel switch, and the McDATA Sphereon and Intrepid switches.

Depending on the firmware levels of these devices, Storage Control supports native interfaces to the device, which simplifies configuration setup and improves device management reliability. After being discovered, these devices will be listed as being managed by Storage Manager in Flex System Manager.

To open Storage Management choose **Home** → **Storage Control** in the horizontal menu in IBM FSM Explorer as shown in Figure 3-82.

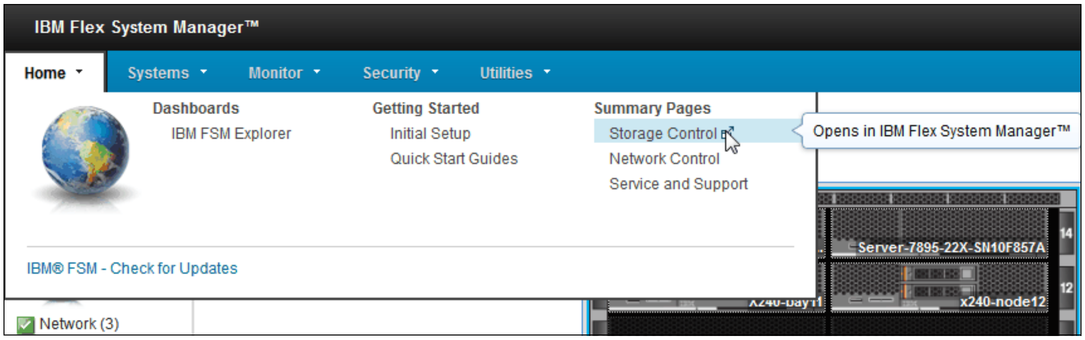


Figure 3-82 IBM FSM Explorer - Storage Control

The Storage Management window opens, showing all storage that is available in the managed IBM Flex System(s), as shown in Figure 3-83.

Figure 3-83 FSM Storage Control

You can also manage storage by going to the **Home tab** → **Additional Setup** → **Manage System Storage**, as shown in Figure 3-84 on page 141.

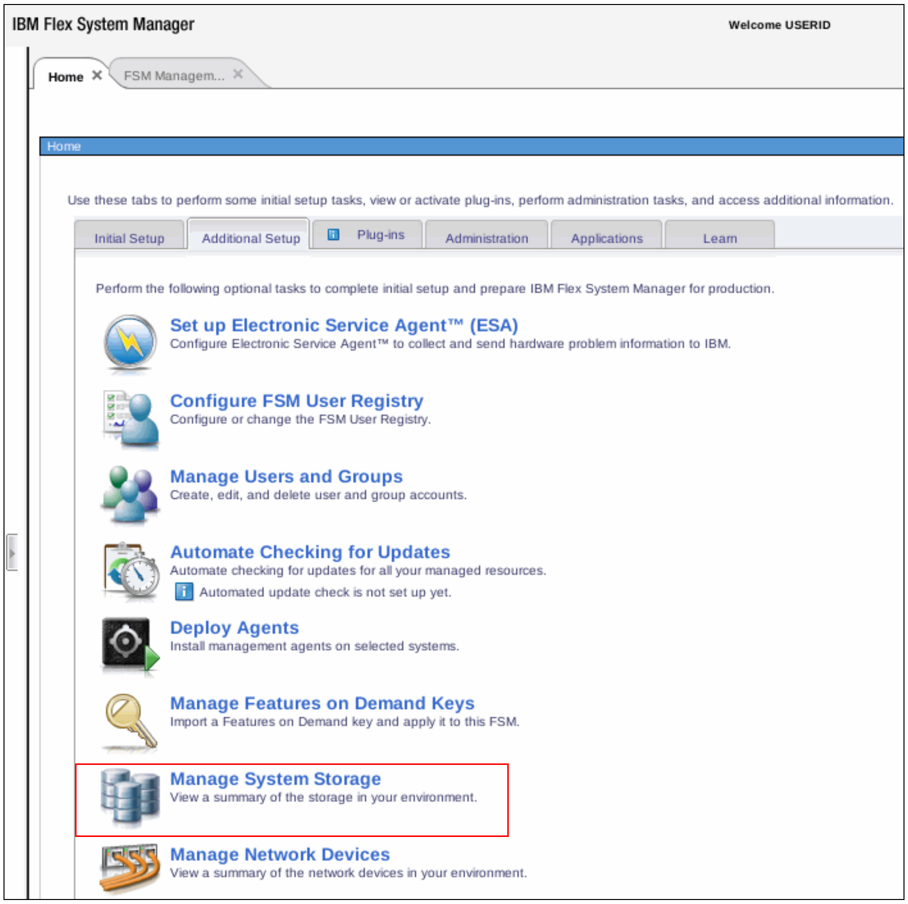


Figure 3-84 FSM - Manage System Storage

To get more information about the storage systems controlled by IBM Flex System Manager and see which actions there can be performed, click on *Network Storage* in the storage control window, as shown in Figure 3-85 on page 142.

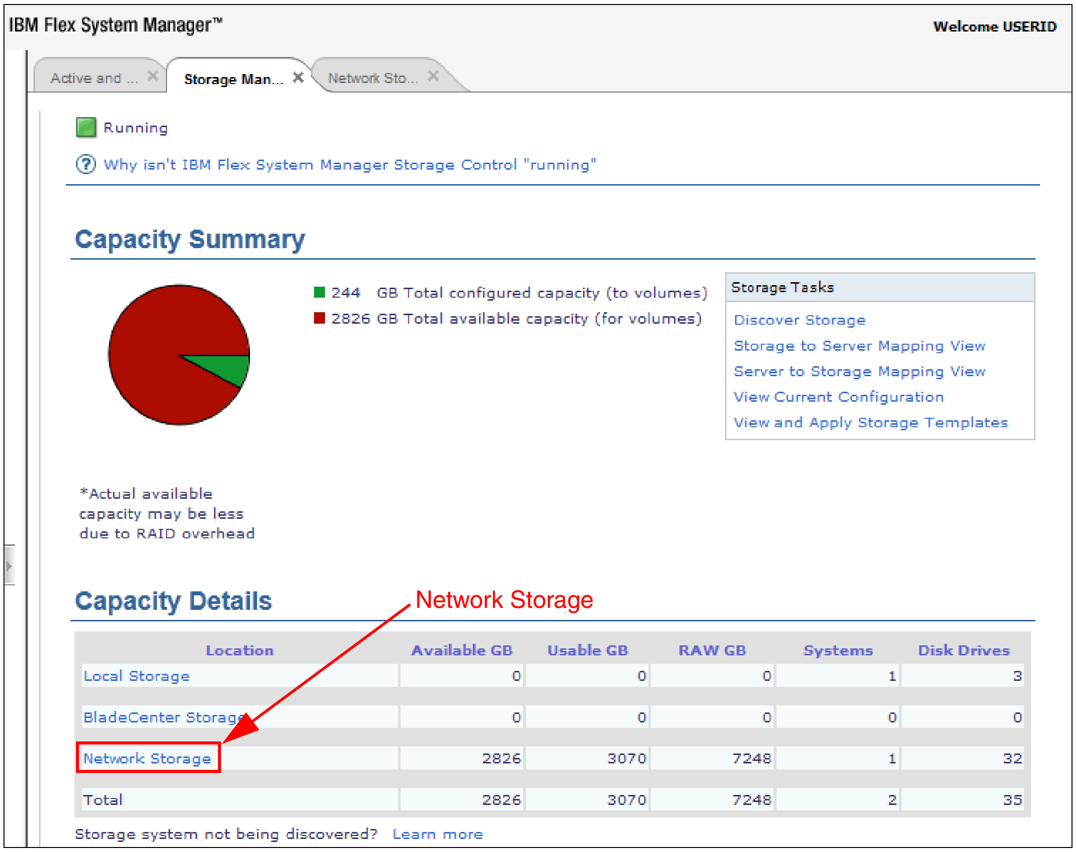


Figure 3-85 FSM - Storage Control - Network storage

This will open a new page showing Network attached storage, as shown in Figure 3-86.

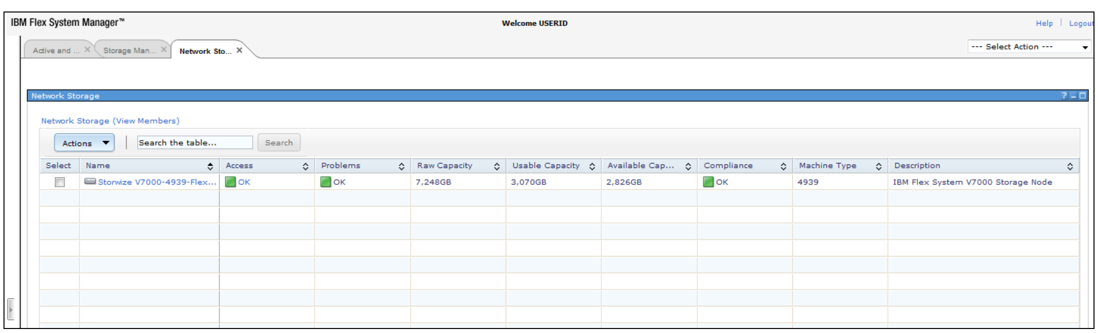


Figure 3-86 FSM - Storage Control - Storage systems

Clicking on the storage system name opens the selected storage systems properties page as shown in Figure 3-87 on page 143.

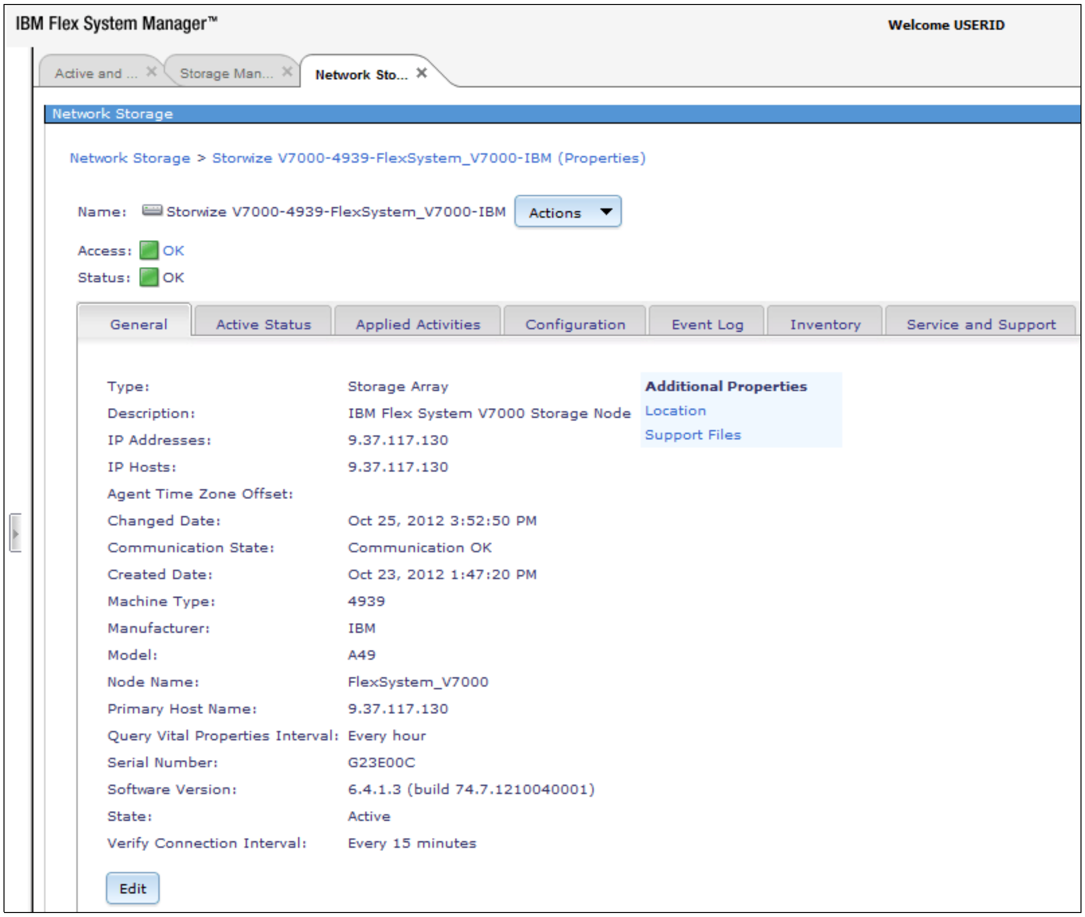


Figure 3-87 FSM - Storage Control - Storage systems

From this window, you are able to see General information, Eventlog and Inventory of the storage system, by clicking on the different tabs.

Choosing the Actions pull down button, will allow you to collect inventory, update storage system firmware and open the IBM Flex System V7000 Storage Node management application, as shown in Figure 3-88 on page 144.

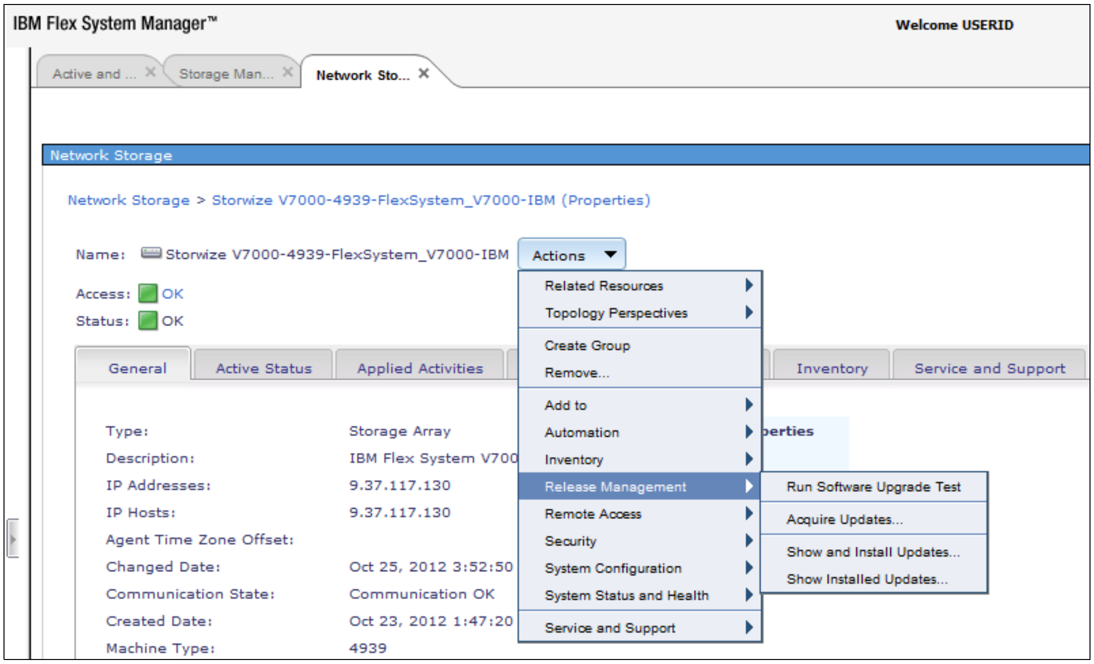


Figure 3-88 FSM - Storage Control - Storage system action menu

Managing External Storage Systems

To manage external storage subsystems in IBM Flex System Manager, perform the following steps:

- From the left slide-out tab, select Resource Explorer, as shown in Figure 3-89.

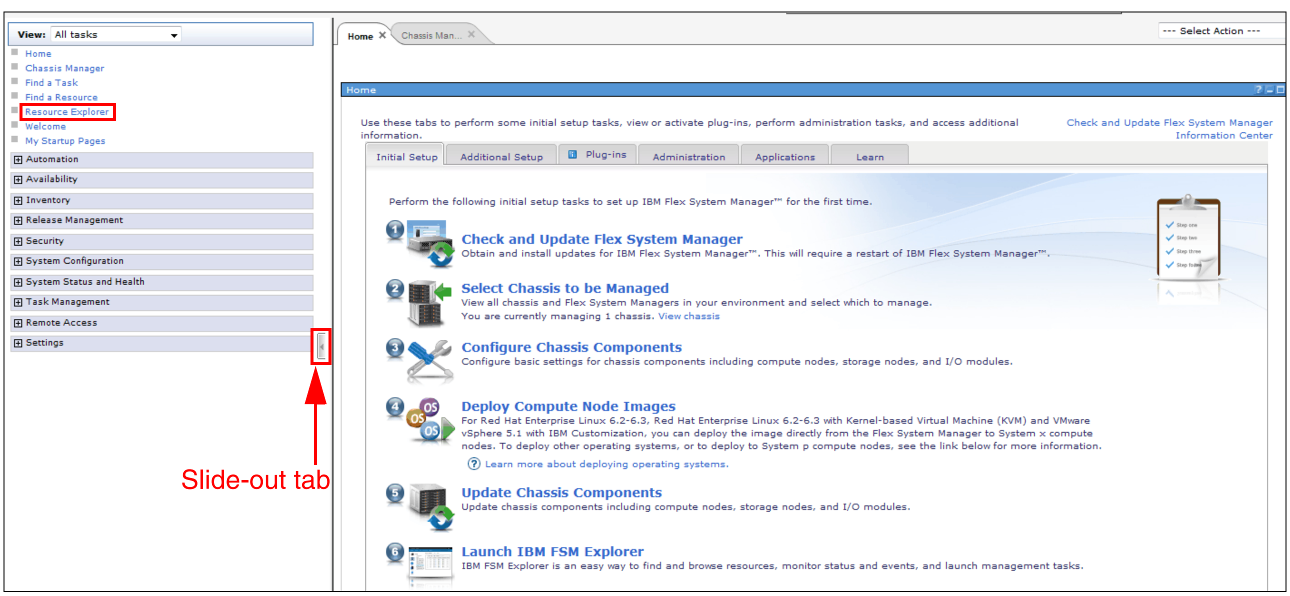


Figure 3-89 Opening Resource Explorer in FSM

- Resource Explorer shows a snapshot of all resources in the IBM Flex System Manager managed by the IBM Flex System Manager, as shown in Figure 3-90 on page 145.

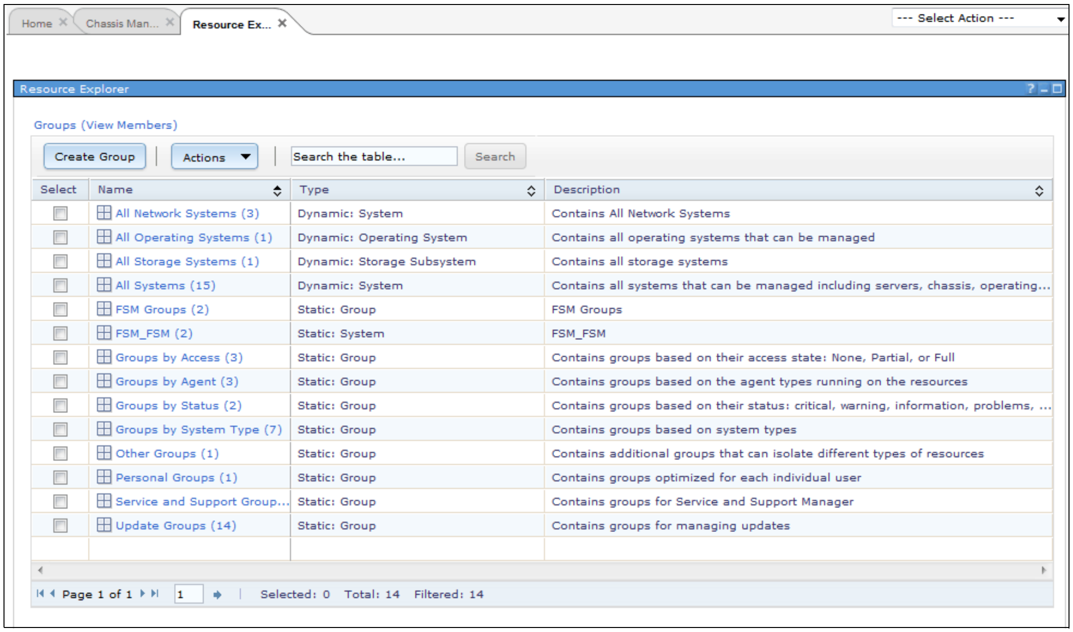


Figure 3-90 Resource Explorer

- Navigate to manage discovered storage, as shown in Figure 3-91.

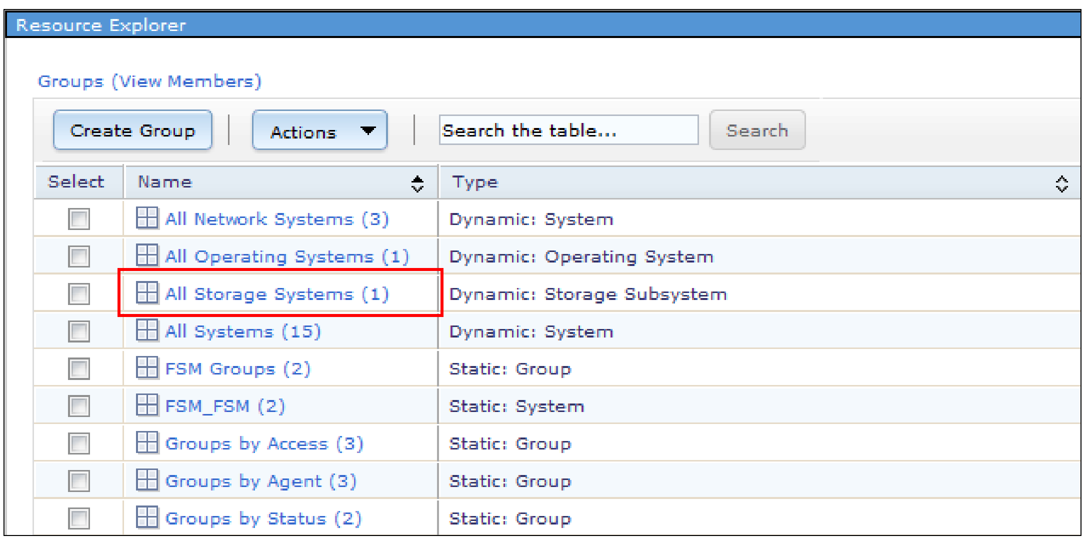


Figure 3-91 Navigate to all storage subsystems

- Select target storage (e.g. IBM Storwize V7000), as shown in Figure 3-92 on page 146.

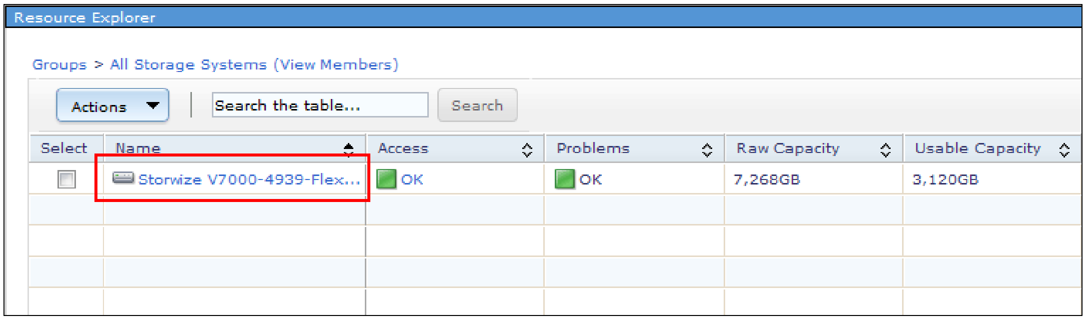


Figure 3-92 Selecting target storage

- Select the tab Inventory to collect the inventory, as shown in Figure 3-93.

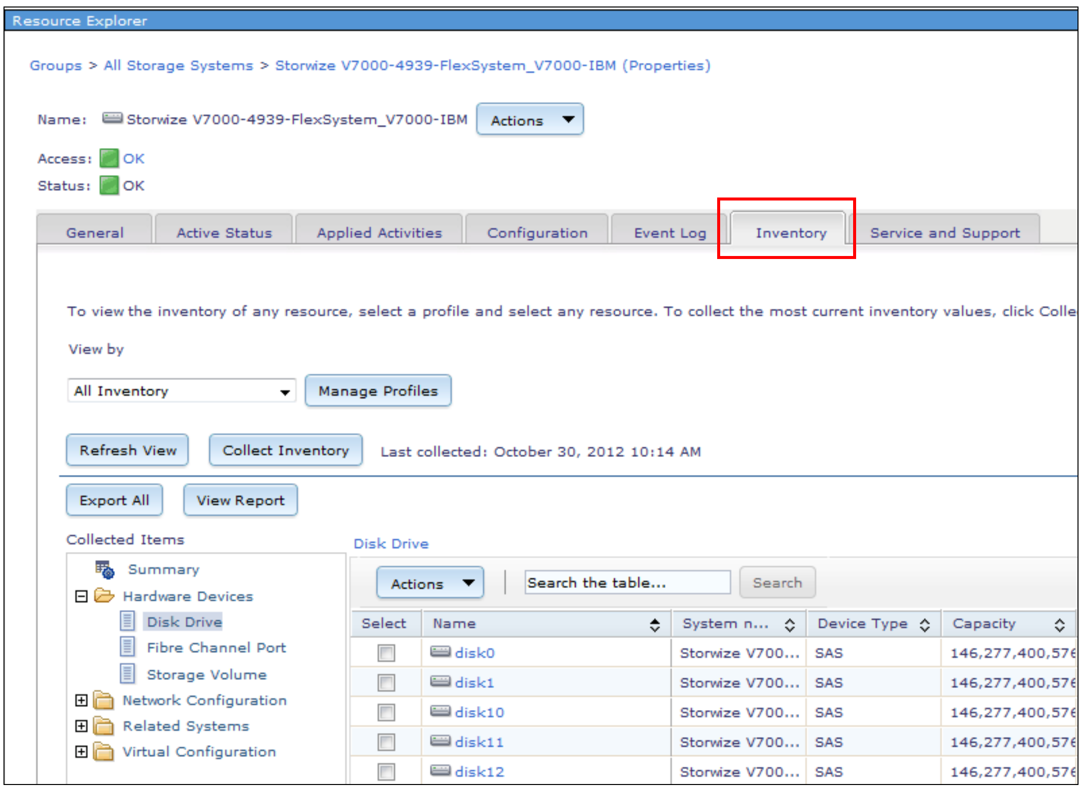


Figure 3-93 Inventory of Storage

- A Task Launch Dialog box is opened to verify the name of the inventory, and time for running it, as shown in Figure 3-94 on page 147.

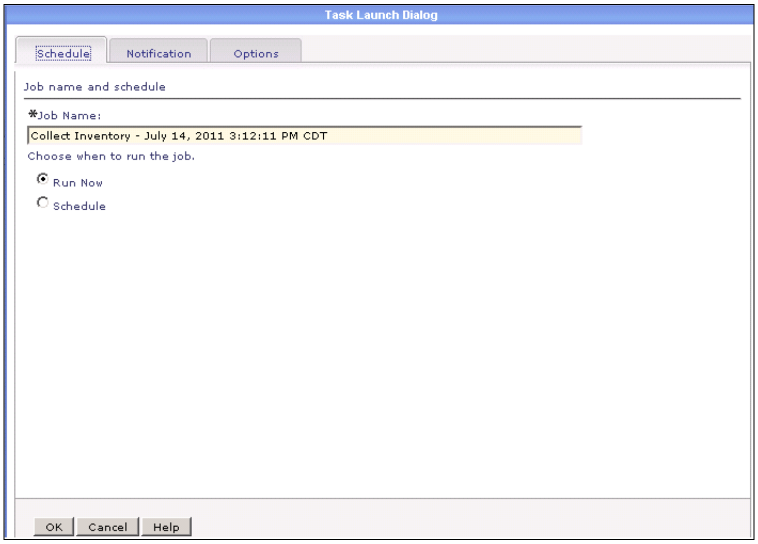


Figure 3-94 Task Launch Dialog

You can use Storage Control to monitor the health of storage subsystems, storage pools, storage volumes, and fabric switches.

Run the inventory against the Storage Control farm to see updated health status. Follow these steps to retrieve the health and alert status for the devices managed by Storage Control.

1. Select **Resource Explorer** → **All Systems**. Select the managed device for which you want to get the latest health and alert status.
2. Discover and then unlock the device.

> **Note:** At this point, you will receive TPC alerts from the farm, but you will not yet receive alerts specific to any devices under this farm.

3. Select the farm that covers the selected device and collect inventory against it.
4. For DS3000, DS4000, and DS5000 TPC devices only: Under the farm, right-click the DSnnnn device, and select Collect inventory.

> **Important:** You must repeat this step each time you want to check for an event on these devices

5. You will now receive TPC alerts for devices under the farm.

In this chapter we have covered management of the IBM Flex System V7000 Storage Node from CMM and FSM. For more information regarding CMM and FSM, refer to *Implementing Systems Management of IBM PureFlex System*, SG24-8060 and IBM InfoCenter:

`http://publib.boulder.ibm.com/infocenter/flexsys/information/index.jsp?topic=%2Fcom.ibm.acc.pureflex.doc%2Fp7een_template_landing.html`

4

IBM Flex System V7000 Storage Node Initial Configuration

In this chapter we discuss the initial configuration steps of the IBM Flex System V7000 Storage Node. Before the actual installation and the initial configuration, proper planning is important. In this chapter, we also present an overview of the planning tasks and refer you to the appropriate documentation as needed to enable you to complete these tasks.

We will discuss the following topics to help you get started, and guide you to a successful install and initial configuration of your of your IBM Flex System V7000 Storage Node system:

- Planning and Installation requirements
- Initial setup with IBM Flex System Manager Node and IBM Flex System Chassis Management Module
- Setup Wizard for initial configuration
- System Management
- Service Assistant

4.1 Planning Overview

We start with a overview of the planning tasks required for the proper installation of IBM Flex System V7000 Storage Node.

4.1.1 Hardware Planning

Proper planning before the actual physical installation of the hardware is required. Below is a overview of the requirements to use to plan your installation:

- Each IBM Flex System V7000 Storage Node enclosure within the IBM Flex System Enterprise Chassis uses 4 bays (double wide and double high). Make sure you have enough space in the chassis for the installation of IBM Flex System V7000 Storage Node control and internal expansion enclosures.
- IBM Flex System V7000 Storage Node control and expansion enclosures are attached in a single redundant SAS chain. Each IBM Flex System V7000 Control Enclosure will support up to nine expansion enclosures total which constitutes an IBM Flex System V7000 Storage Node.
- IBM Flex System V7000 Control Enclosure supports the attachment for up to nine Expansion Enclosures using either IBM Flex System V7000 Expansion Enclosures (max 2), IBM Storwize V7000 Expansion Enclosures (max 9), or a combination of both.
- The IBM Flex System Enterprise Chassis contains 14 bays in total, so there can be at most three IBM Flex System V7000 Storage Node enclosures (a control enclosure and up to two IBM Flex System V7000 Expansion Enclosures). Any additional expansion enclosures would have to be IBM Storwize V7000 expansion enclosures attached externally to the same single SAS chain.

  > **Note:** Installing a Flex System V7000 Enclosure within an Enterprise Chassis requires the removal of some of the shelves within the chassis which would normally support the Compute Nodes.

  Install the hardware as described in IBM Flex System V7000 Storage Node Installation Guide (GC27-4203)

  For information about planning the IBM IBM Flex System V7000 Storage Node environment, see the following address:

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_product_page.html

  For the latest maximum configuration support information, see the IBM System Storage Flex System V7000 Storage Node website:

  http://www.ibm.com/support/entry/portal/overview/hardware/puresystems/pureflex_system/storage_node/flex_system_v7000
- There needs to be two power outlets from redundant power sources in the rack for the two power cords included with each enclosure. 2 - 20 outlets are required depending on the number of enclosures to install. The power cords connectors are IEC320 C13/C14 standard connections.
- There are four optional configurations that can exist for host connections to the IBM Flex System V7000 Storage Node through network adapters. See Chapter 11, "SAN Connections and Configuration" on page 431 for detailed examples of how to configure the host connections.

- You should have a minimum of three IP addresses for initial system configuration. These IP addresses can be of either the IPv4 or IPv6 standards. One for management and two for service access to each of the control canisters.
- A minimum of one and up to four IP addresses are needed if iSCSI attached hosts access volumes from the IBM Flex System V7000 Storage Node. These are in addition to the three IP addresses needed for the management of the system.

4.1.2 SAN configuration planning

The SAN fabric is an area of the network that contains routers and switches. The IBM Flex System V7000 Storage Node has connections for up to eight 8Gbps fiber channel connections per control enclosure. A SAN is configured into a number of zones. A device using the SAN can communicate only with devices that are included in the same zones that it is in. A Flex System V7000 Storage Node can have several distinct types of zones: a system zone, host zones, and disk zones. The intersystem zone is optional. Some operating systems cannot tolerate other operating systems in the same host zone, although you might have more than one host type in the SAN fabric. For example, you can have a SAN that contains one host that runs on an IBM AIX operating system and another host that runs on a Microsoft Windows operating system.

After the IBM Flex System V7000 Storage Node, hosts, and optional external storage systems are connected to the SAN fabrics, zoning needs to be implemented. See Chapter 11, “SAN Connections and Configuration” on page 431 for more information.

Note: All communication between the canisters in the same control enclosure is performed either through the SAN or via PCI Express (PCIe) link. All configuration and service commands are sent to the system through an Ethernet network.

The recommended SAN configuration is composed of a minimum of two fabrics with all IBM Flex System V7000 Storage Node ports, host ports, and ports from external storage systems which are to be virtualized divided equally between the two fabrics for redundancy in the event one of the fabrics should go offline (either planned or unplanned).

Degraded performance can occur when you have a SAN configuration in which any single component might fail and connectivity between the devices within the SAN is maintained. Splitting the SAN into two independent counterpart SANs achieves this normally. Because of the requirement for high availability, IBM Flex System V7000 Storage Node generally is installed into a redundant SAN, however, a non-redundant SAN is supported.

4.1.3 LAN configuration planning

The Enterprise Chassis provides the following services to the IBM Flex System V7000 Storage Node:

- 1 Gbps management network
- 4 host attachment networks
  - 10 Gbps Ethernet
  - 8 Gbps or 16 Gbps Fibre Channel Service IP address

4.1.4 Management IP address considerations

The IBM Flex System V7000 Storage Node management IP address provides access to the system management interfaces, including the GUI and CLI. The management IP address is

also used to access remote services like authentication servers, NTP, SNMP, SMTP, and Syslog systems, if configured.

Management IP address is configured during first-time setup of your IBM Flex System V7000 Storage Node. See 4.3, “IBM Flex System V7000 Storage Node Setup Wizard” on page 161.

The management IP address is associated with one of the control canisters in the IBM Flex System V7000 Control Enclosure and that control canister becomes the *configuration node*. Should this node go offline, either planned or unplanned, the management IP address does a failover to the surviving partner control canister’s Ethernet port and it becomes the *configuration node*.

Note: IP addresses that are used by hosts to access the system over an Ethernet connection are different from management IP addresses.

Following rules should be observed when configuring a management IP for your IBM Flex System V7000 Storage Node:

- Ethernet port 1 (1 Gbps) must be configured with an IPv4 or IPv6 system address.
- Ethernet port 2 (1 Gbps) can optionally be configured with a second system address.
- A maximum of one IPv4 address and one IPv6 address can be configured on each of Ethernet ports 1 and 2 for system addresses.
- To ensure system IP failover operations, Ethernet port 1 on all nodes must be connected to the same subnet. The system IP address can failover to any control node in the system.
- If Ethernet port 2 (1 Gbps) is configured with a system IP address, Ethernet port 2 on all nodes must also be connected to the same subnet. However, the subnet for Ethernet port 2 does not have to be the same as Ethernet port 1.
- Management addresses cannot be configured on ports 3 or 4 (10 Gbps), if present.

4.1.5 Service IP address considerations

The Service Assistant is a web application that helps troubleshoot and resolve problems in a Flex System V7000 Storage Node enclosure. Service Assistant is connected to on each IBM Flex System V7000 Control Enclosure node canister through the Service IP address.The node canister cannot be active as part of a system while it is in service state.

The Ethernet port 1 on each node canister is used for system management, and also used for service access when required. In normal operation, the Service IP addresses are not needed However, in the event of a node canister problem, it might be necessary for service personnel to log on to the node to perform service actions.

Note: The password is required to access the system through the Service IP address. The authentication of superuser is always local; therefore, this user ID can be used even when a remote authentication server that is used for other users is not available.

To establish connectivity, the management node attempts to use Dynamic Host Configuration Protocol (DHCP) to acquire its initial IP address for the Chassis Management Module (CMM) Ethernet port. If DHCP is not installed or is enabled and fails, the management node uses the static IP address.

4.1.6 Management interface planning

IBM PureFlex Systems are pre-configured, pre-integrated infrastructure systems with different management options:

- Chassis Management Module (CMM)
- Flex System Manager (FSM)

For information and the actual implementation of the management modules, refer to *Implementing Systems Management of IBM PureFlex System*, SG24-8060

Chassis Management Module (CMM)

The IBM Flex System Enterprise Chassis Management Module (CMM) is a hot-swap module that configures and manages all installed chassis components. The chassis comes with one Chassis Management Module in the CMM bays.

The IBM Flex System Chassis Management Module auto-detects and manages the hardware of the entire Flex chassis and provide centralized system monitoring and management.

> **Note:** The IBM Flex System Chassis Management Module provides control and information on the storage hardware, while the IBM Flex System Manager in addition provides configuration functionality.

For more details, see Chapter 3, “Systems Management” on page 81.

The IBM Flex System Manager Node (FSM)

The Flex System Manager (FSM) is an advanced chassis management appliance which provides the ability to manage up to four enterprise chassis. It is designed with a specialized compute node, storage, networking, and provides physical and virtual management capabilities.

The FSM appliance is based on an Intel-based x86 compute node that comes with preloaded management software. The FSM software contains a fully operational management stack for managing multiple chassis from a single focal-point for support of all assets and resources in IBM PureFlex Systems handling the complete range of management tasks. The following is a list of the tasks that can be performed from by the FSM:

- Service requests
- Incident handling and problem resolution
- Handling specialized management tasks in connection with:
  - Security
  - Performance
  - Resource allocation
  - Application version control

4.2 Initial setup for IBM Flex System V7000 Storage Node

When a new IBM Flex System V7000 Storage Node control enclosure is installed in the Flex System chassis for the first-time you must create and configure it into *clustered system*. The following procedure will guide you through the necessary steps to set up your IBM Flex System V7000 Storage Node system for the first time in this configuration.

There are two methods that can be used for the initial setup of the IBM Flex System V7000 Storage Node. The method used depends upon the configuration of the IBM Flex System.

- If the IBM Flex System has an installed and configured IBM Flex System Manager, then it should be used to set up the IBM Flex System V7000 Storage Node.
- If the IBM Flex System does not have an IBM Flex System Manager (FSM) installed and configured, use the Chassis Management Module to set up the IBM Flex System V7000 Storage Node.

Before starting

- You should have installed the IBM Flex System V7000 Storage Node enclosures in the rack, connected all cables, powered the system on, and checked the LED status for the system. When you have completed these steps, you can continue with the configuration of the IBM Flex System V7000 Storage Node.
- To complete this procedure, use a supported web browser. To verify that you are using a supported web browser from the following website:

  http://www.ibm.com/support/entry/portal/overview/hardware/puresystems/pureflex_system/storage_node/flex_system_v7000
- Decide whether you are going to use an DHCP server to assign the node service IP address or will assign static IP addresses.

  If using an DHCP server make sure the IBM Flex System V7000 Storage Node has access to the DHCP server and there are at least two DHCP addresses available on the subnet for their use. These are the addresses that will be used in the initial installation.

  If assigning static service IP addresses, make sure that you also know the management IP address that you want to assign to the IBM Flex System V7000 Storage Node. This is the address that will be used to access the IBM Flex System V7000 Storage Node cluster for management after installation is complete.
- If you are using the Chassis Management Module (CMM) to setup your system, ensure that the chassis the IBM Flex System V7000 Storage Node is installed in is being managed by the CMM.

  If you are planning to use the CMM to setup your system, verify that you are using a CMM that supports the IBM Flex System V7000 Storage Nodes from the following website: http://www.ibm.com/support/entry/portal/overview/hardware/puresystems/pureflex_system/storage_node/flex_system_v7000
- If you are using the Flex System Manager (FSM) to setup your system, you must have an installed and pre-configured FSM. Verify that you are using a supported FSM for IBM Flex System V7000 Storage Nodes from the following website: http://www.ibm.com/support/entry/portal/overview/hardware/puresystems/pureflex_system/storage_node/flex_system_v7000

Note: USB key is not required for first-time setup or normal operations. However, it may be required for resetting password, service IP, and to check the status of an inaccessible canister if directed by IBM Service and Support.

4.2.1 Using FSM for initial setup

Now that you have installed the IBM Flex System V7000 Storage Node in the chassis with supported level of code you can create and configure a clustered system using the FSM. The following procedure will guide you through the necessary steps when using FSM web user interface.

1. Open a web browser and point it to the IP address of the FSM and login. The following menu screen (Figure 4-1) appears giving you a number of selections.

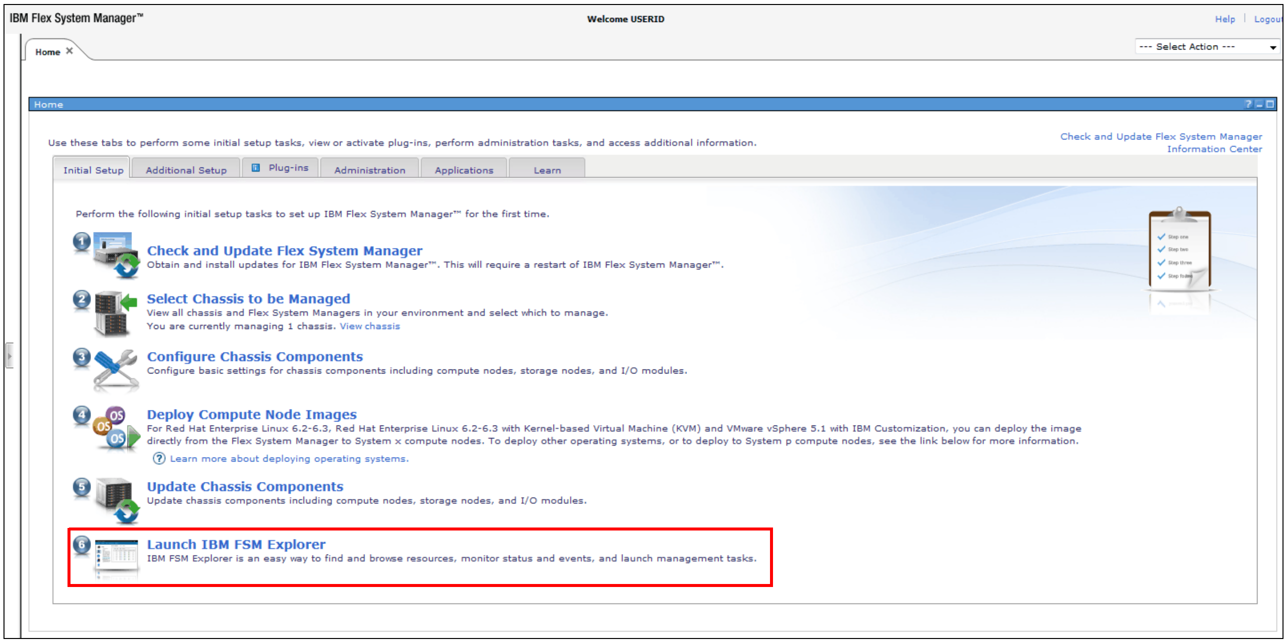


Figure 4-1 Select and launch the chassis in the Chassis Manage

Select Launch IBM FSM Explorer from above menu list.

Notice a new browser tab is opened, which allows you to select the applicable enclosure from the Chassis Map as shown in Figure 4-2.

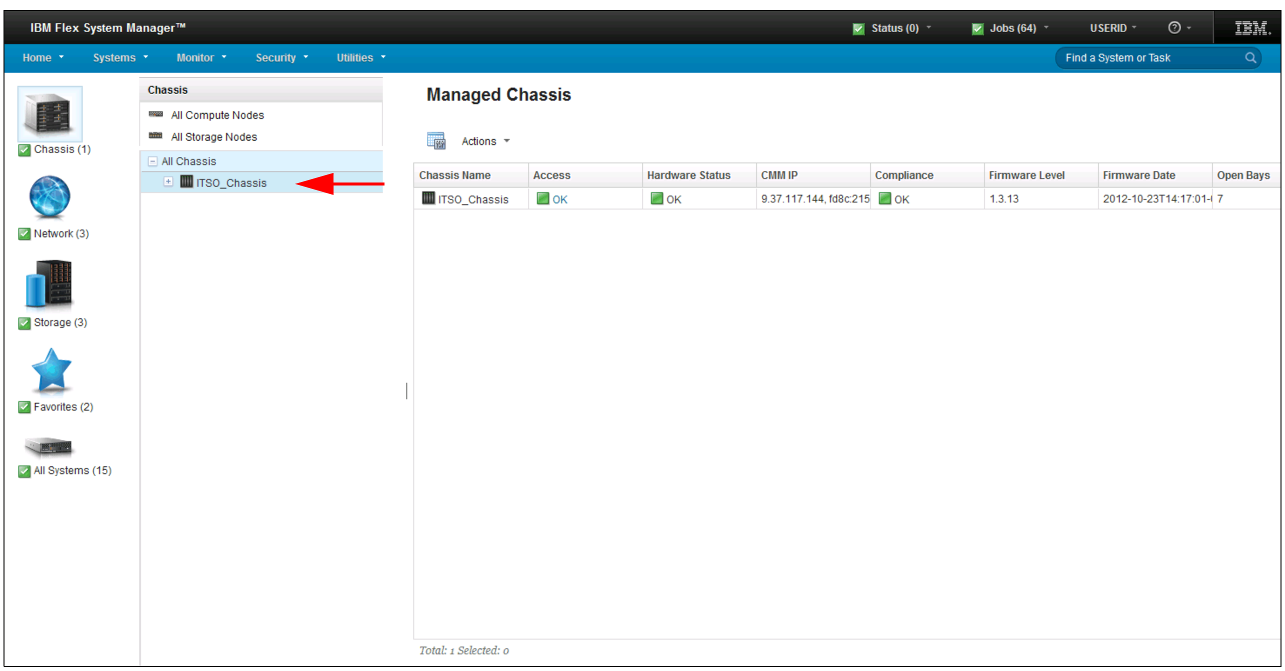


Figure 4-2 Select and launch the chassis in the Chassis Manage

2. In the Chassis Manager, select the applicable chassis which will launch the Chassis Map for that chassis, as shown in Figure 4-3 on page 156.

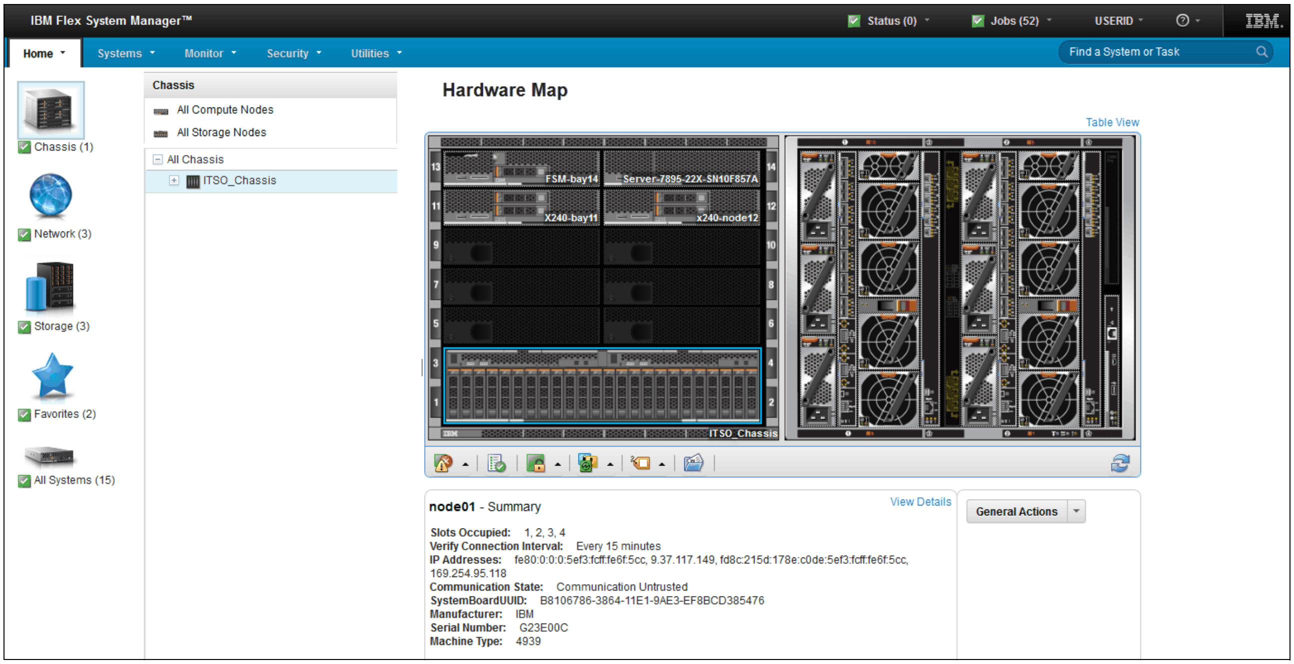


Figure 4-3 IBM Flex System Manager - Hardware Map

3. From the Chassis Manager page in the management software web interface, click the chassis name that contains your applicable storage node as shown in Figure 4-4 on page 157 (A) and continue to Step 4 on page 157. If the chassis is not displayed, add the chassis by performing the following steps:
   – Select **General Actions** → **Add or Remove a chassis**, as shown in Figure 4-4 on page 157 (B).
   – Select the chassis you want to manage
   – Click Manage. The Management Domain page displays.
   – Click Manage.
   – Enter the user ID and password for the CMM of the chassis that you want to manage.
   – Wait while the status is processed. After the status is processed, the status changes to Success, and the message changes to Managed.
   – Click Done.

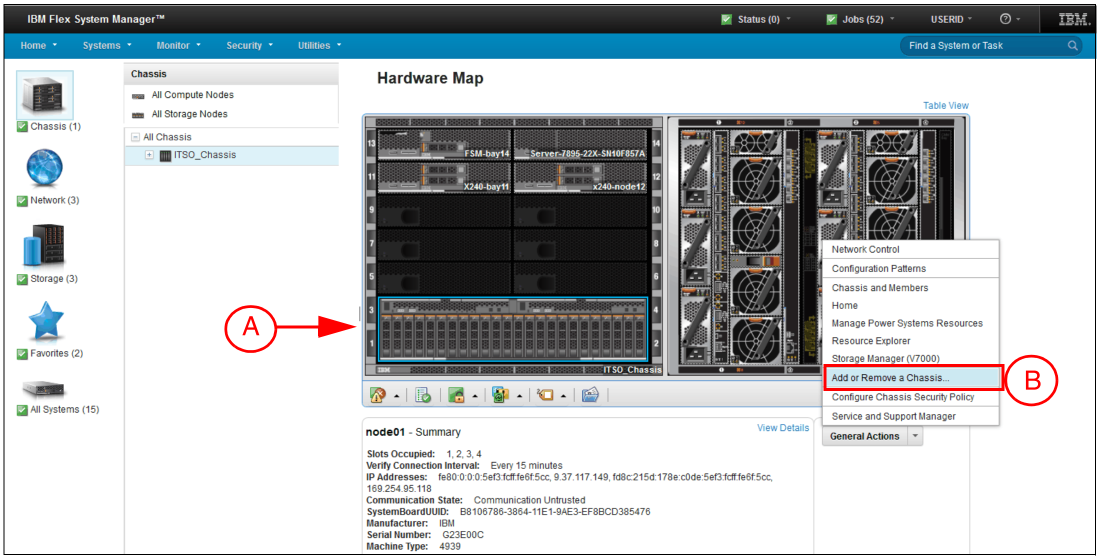


Figure 4-4 Add a chassis in to Chassis Manager

4. Navigate to General Actions and click Launch IBM Flex System V7000 Storage Node Manage (V7000) as shown in Figure 4-5 to start the Initial Setup wizard. If you do not see this option, examine the Service IP setting for the storage node. See to 4.5.1, “Change Service IP address” on page 170, for details.

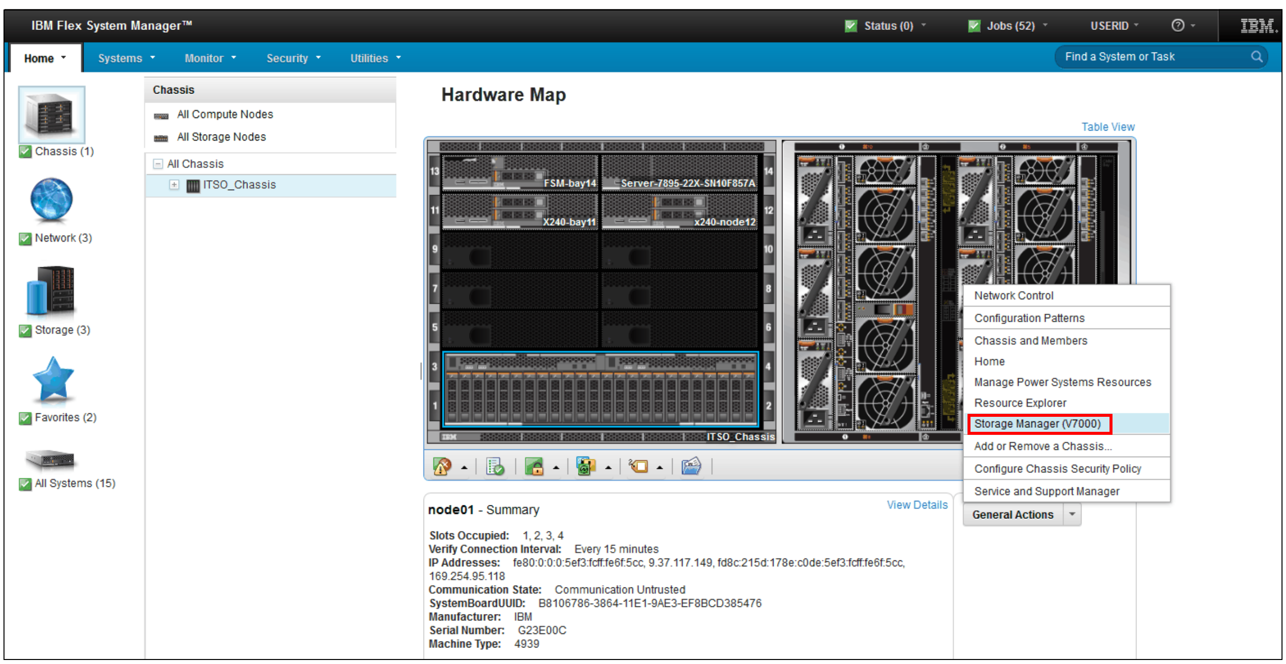


Figure 4-5 Launch Storage Manager (V7000)

5. Next window is a welcome window from IBM Flex System V7000 Storage Node interface, asking to either create a new system (cluster) or add to an existing system, as shown in Figure 4-6 on page 158

Figure 4-6 IBM Flex System V7000 Storage Node first time setup welcome window

6. If you choose to create a new system, it will ask for network information, as shown in Figure 4-7. Select whether you are using an IPv4 or IPv6 management IP address and type in the address (you can use either DHCP or static address that was assigned). The subnet mask and gateway will already have defaults listed, which you can edit.

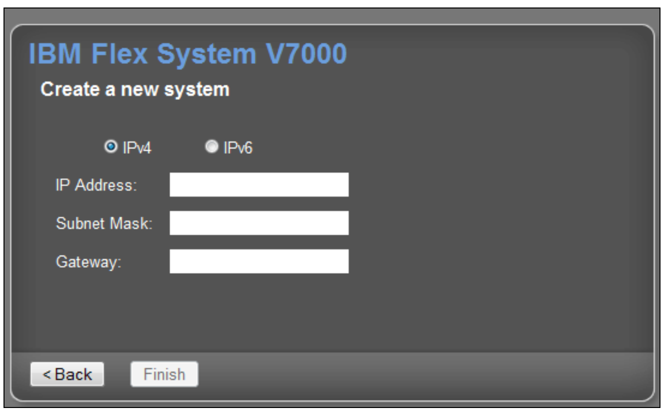


Figure 4-7 Create new Storage Cluster

7. Click Finish to set the management IP address for the system. System initialization begins and might take several minutes to complete.

   When system initialization is complete, System Setup is launched automatically. The setup wizard will take you through the steps to configure basic system settings such as time and date, system name, and hardware detection and verification.

   For more details on the rest of the setup wizard steps, please see 4.3, "IBM Flex System V7000 Storage Node Setup Wizard" on page 161.

4.2.2 Using CMM for Initial setup

When the first control enclosure for a new IBM Flex System V7000 is installed in the Flex System chassis, you must create and configure a system. This procedure is used to guide you through the necessary steps using the CMM for the first time.

Before you begin

> **Attention:** If the control enclosure you have installed in the chassis is being added to an existing clustered system, do not use this procedure. Go to Adding another control enclosure into an existing system.

- You need to be sure your IBM Flex System V7000 Storage Node has access to the DHCP server and there are at least two DHCP addresses available on your subnet. These are the addresses that will be used in the initial installation. If DHCP addresses are unavailable, the node will use a default service IP address based on the slot location. This can be changed to a static address using the CMM Component IP Configuration page.
- When a new control enclosure is installed in the IBM Flex System Enterprise Chassis, the CMM automatically recognizes the new hardware. Open the CMM web user interface and navigate to the chassis map. In the chassis map displayed by the CMM, you will see the new Flex System V7000 Storage Node.

Follow the below procedure for the initial setup from the CMM interface.

1. Log in to the CMM and navigate to the chassis map. In the chassis map displayed by the CMM, you will see all the IBM Flex System V7000 Storage Node enclosures installed in the chassis.

   > **Note:** If you have installed more than one control enclosure in your IBM Flex System V7000 Storage Node system, choose a control enclosure to be used to initialize the system. It doesn't matter which control enclosure you choose. You cannot initialize the system from an expansion enclosure.

2. Select the Actions menu and click Launch Storage Node Console. The Action menu can be found by right-clicking on the left canister in the graphical representation of the storage node, as shown in Figure 4-8.

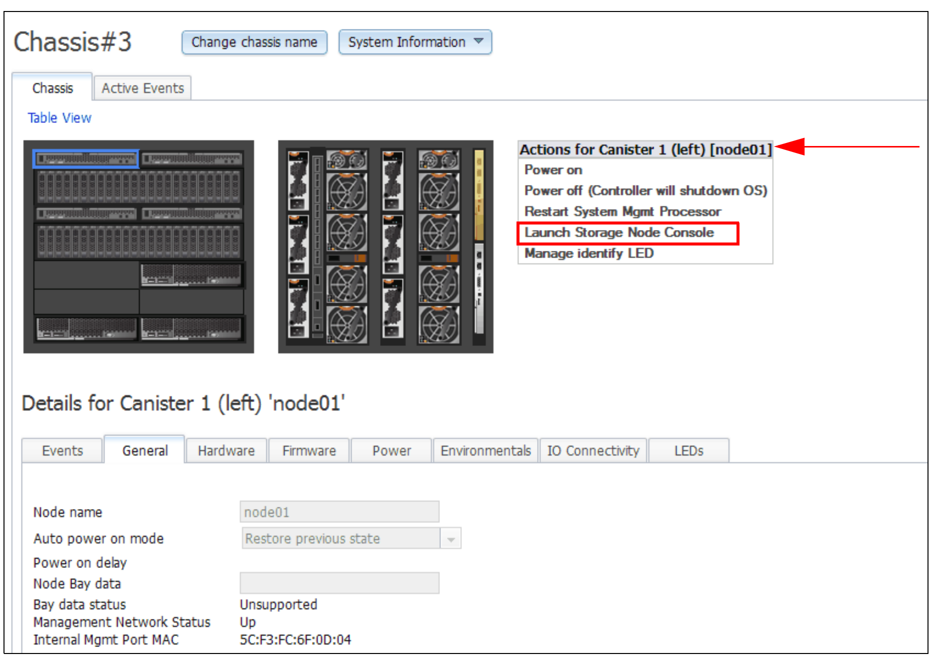


Figure 4-8 CMM - Select Launch Storage Node Console

3. A browser window displays the IBM Flex System V7000 Storage Node system startup screen. The Service IP address displayed will be one that has been allocated by your DHCP server as shown in Figure 4-9.

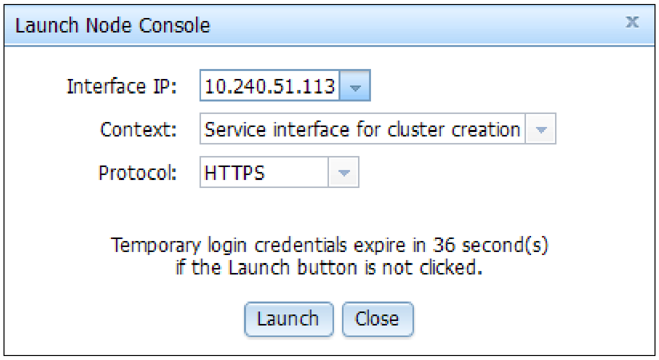


Figure 4-9 CMM - Launch Node Console

 – If a default Service IP address in the range between 192.168.70.131 and 192.168.70.144 is displayed, the canister has not obtained a DHCP address. This default service IP address can be changed to a static address using the CMM Component IP Configuration page. Refer to the *Changing the service IP address of a node canister topic* in the IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide PDF on the CD provided with your system or in the IBM Flex System V7000 Storage Node Information Center.
 – If you have problems connecting to the service address, see the Cannot connect to the service assistant topic in the IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide PDF on the CD or in the IBM Flex System V7000 Storage Node Information Center.

After clicking Launch in Figure 4-9, you will be have an option to Create a new system or Add to an existing system as shown in Figure 4-10. We select Create a new system then click Next.

Figure 4-10 IBM Flex System V7000 Storage Node first time setup welcome window

4. Select whether you are using an IPv4 or IPv6 management IP address and type in the address (you can use DHCP or statically assigned one). The subnet mask and gateway will already have defaults listed, which you can edit, as shown in Figure 4-11 on page 161.

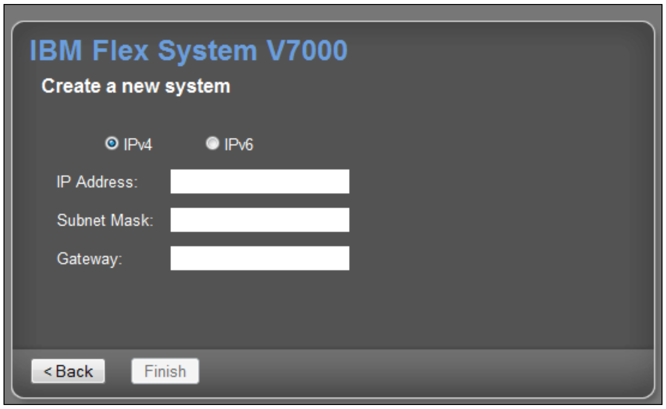


Figure 4-11 Create new Storage Cluster

5. Click Finish Figure 4-11 to set the management IP address for the system. System initialization begins and might take several minutes to complete.

   When system initialization is complete, System Set Up is launched automatically. The setup wizard will take you through the steps to configure basic system settings such as time and date, system name, and hardware detection and verification.

   For more details on the rest of the setup wizard steps, please see 4.3, "IBM Flex System V7000 Storage Node Setup Wizard" on page 161

4.3 IBM Flex System V7000 Storage Node Setup Wizard

After the initial configuration using the FSM or CMM, we can continue with the setup wizard for the rest of the configuration of IBM Flex System V7000 Storage Node.

After clicking Finish on Figure 4-11, on the next screen, enter the system name and a new superuser password. At this point you have the option to use the setup wizard to configure more settings such as notifications and storage configuration. If you are not ready to complete more configuration steps now, use the configuration tasks in the management GUI to complete the configuration at a later time.

When you complete the Initial Setup wizard, return to the FSM or CMM console.

> **Note:** It is recommended that you upgrade to the most current level of software after installing the Flex System V7000 Storage Node. The management GUI can be used to apply software updates. Refer to the IBM Flex System V7000 Storage Node support website for the latest information about software upgrades.

1. After the initial configuration described in 4.2, "Initial setup for IBM Flex System V7000 Storage Node" on page 153 is complete, the IBM Flex System V7000 Storage Node Welcome window opens (Figure 4-12 on page 162).

Figure 4-12 IBM Flex System V7000 Storage Node Welcome window

> **Note:** During the initial setup of the Flex System V7000, the installation wizard asks for various information that you should have available during the installation process. If you do not have this information or choose not to configure some of the items at this time, you can configure them later through the GUI.

2. You have to Read and accept the license agreement to be able to continue as shown in Figure 4-13. Click Next after accepting the licence agreement.

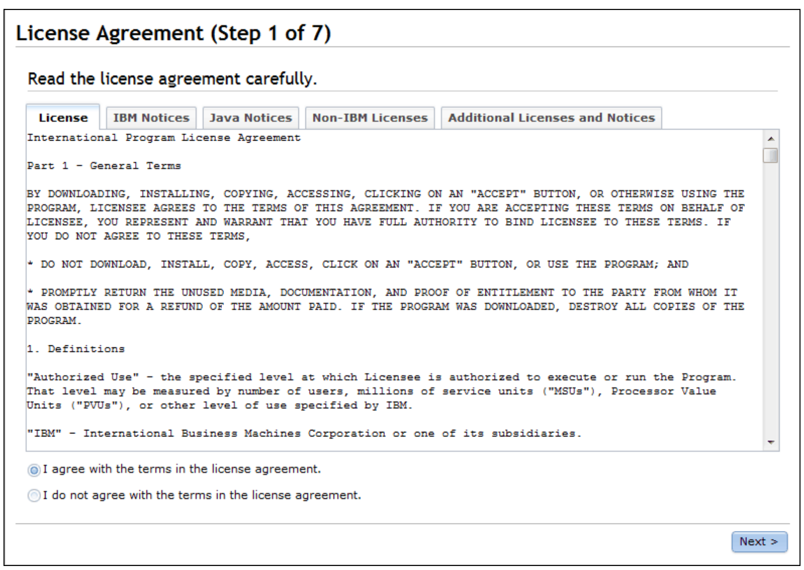


Figure 4-13 Setup wizard - License Agreement

3. Set up the System Name and the Superuser Password and click Next as shown in Figure 4-14 on page 163.

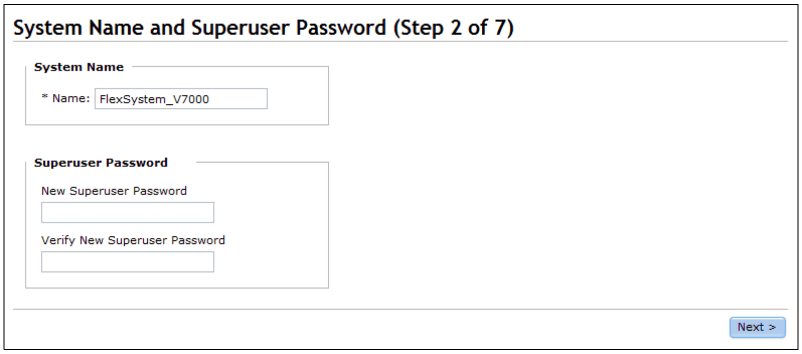


Figure 4-14 Setup wizard - Set system name and superuser password

4. Set up the system date and time and click Next as shown in Figure 4-15.

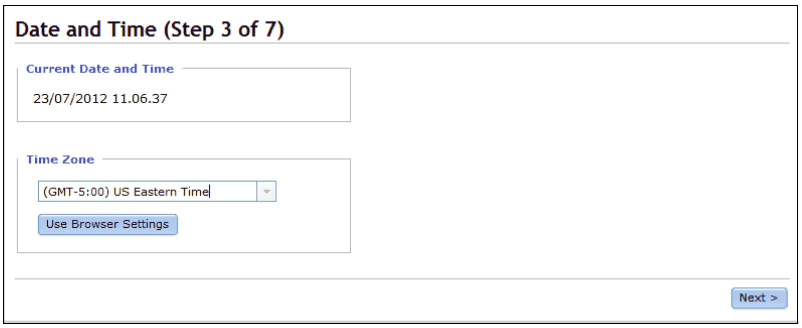


Figure 4-15 Setup wizard - Set date and time

5. Optionally, you can type in System licenses and click Next as shown in Figure 4-16. The System Licenses include External Virtualization Limit, Remote-Copy Limit, and Real-time Compression Limit. The virtualization license for all directly attached expansion enclosures are already included in the system license and should not be added here.

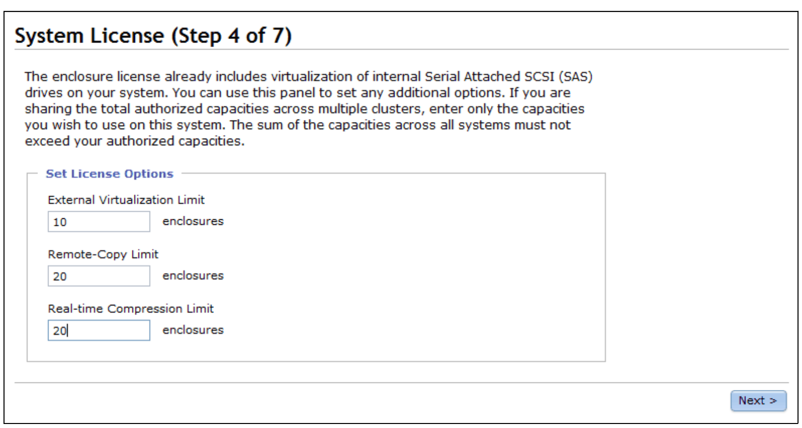


Figure 4-16 System license

6. Configure support notifications as shown in Figure 4-17 on page 164.

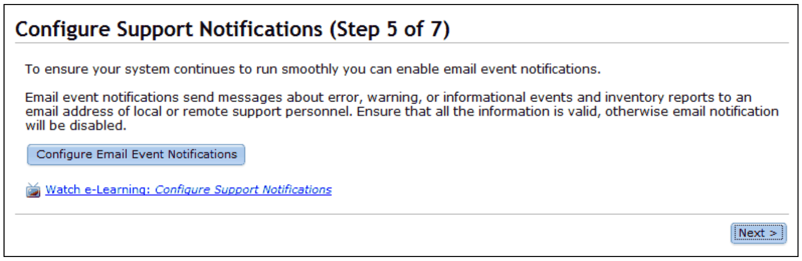


Figure 4-17 Configure support notifications

7. Define company contact info as shown in Figure 4-18.

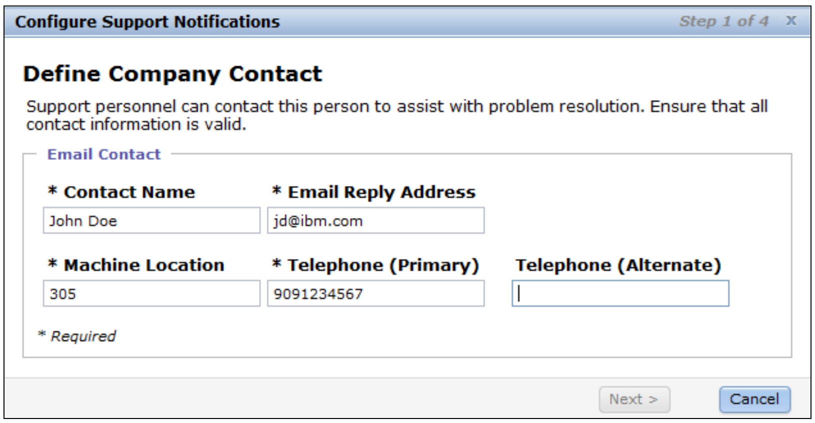


Figure 4-18 Define company contact

8. Verify that all hardware has been detected by the system correctly as shown in Figure 4-19.

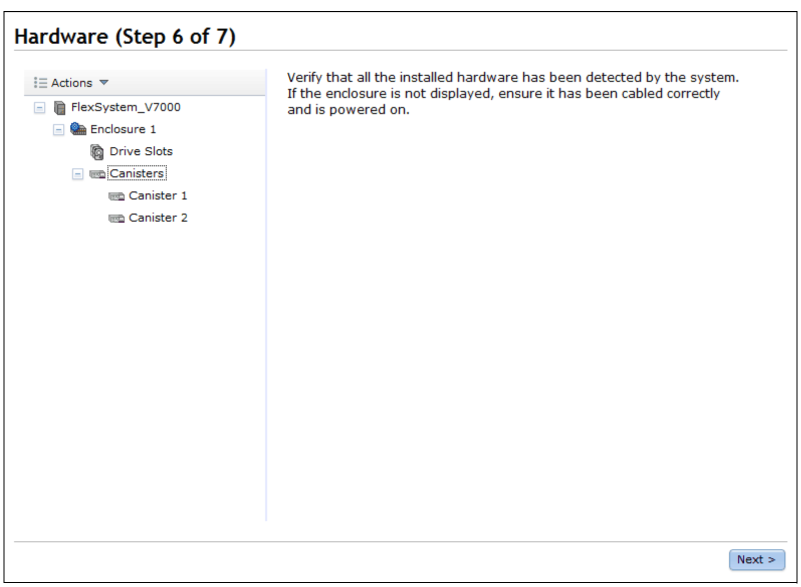


Figure 4-19 Verify hardware

9. Select ‘Yes’ to automatically configure internal storage now if preferred (Figure 4-20 on page 165).

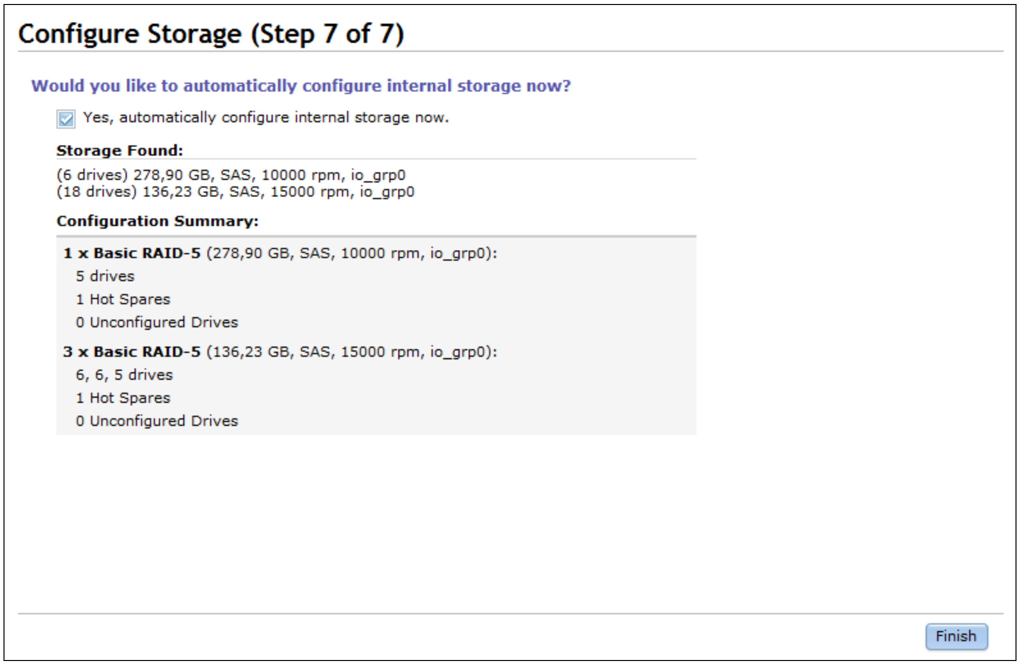


Figure 4-20 Select 'Yes' to automatically configure internal storage now

10. Click Finish to complete Setup wizard task and log in to IBM Flex System V7000 Storage Node as shown in Figure 4-21. You login as superuser with your newly defined password. If you have note changed the password the default is "passw0rd".

Figure 4-21 Setup wizard task complete

11. After successful login, the IBM Flex System V7000 Storage Node Home Overview window is shown in Figure 4-22 on page 166.

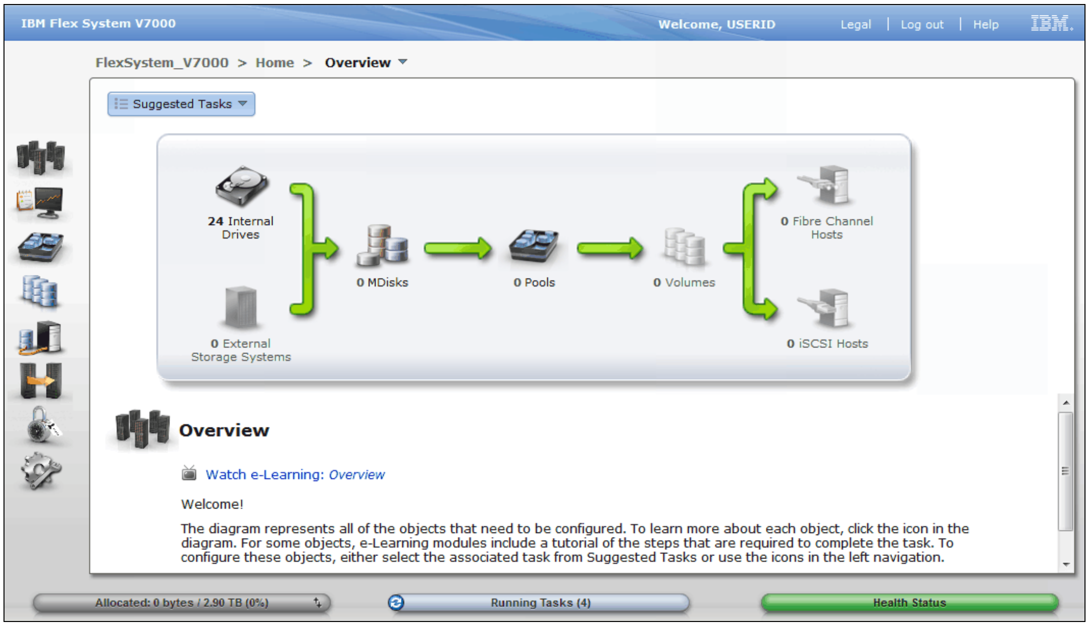


Figure 4-22 IBM Flex System V7000 Storage Node Home Overview window

12. IBM Flex System V7000 Storage Node initial configuration is complete and the cluster is up and running (Figure 4-23).

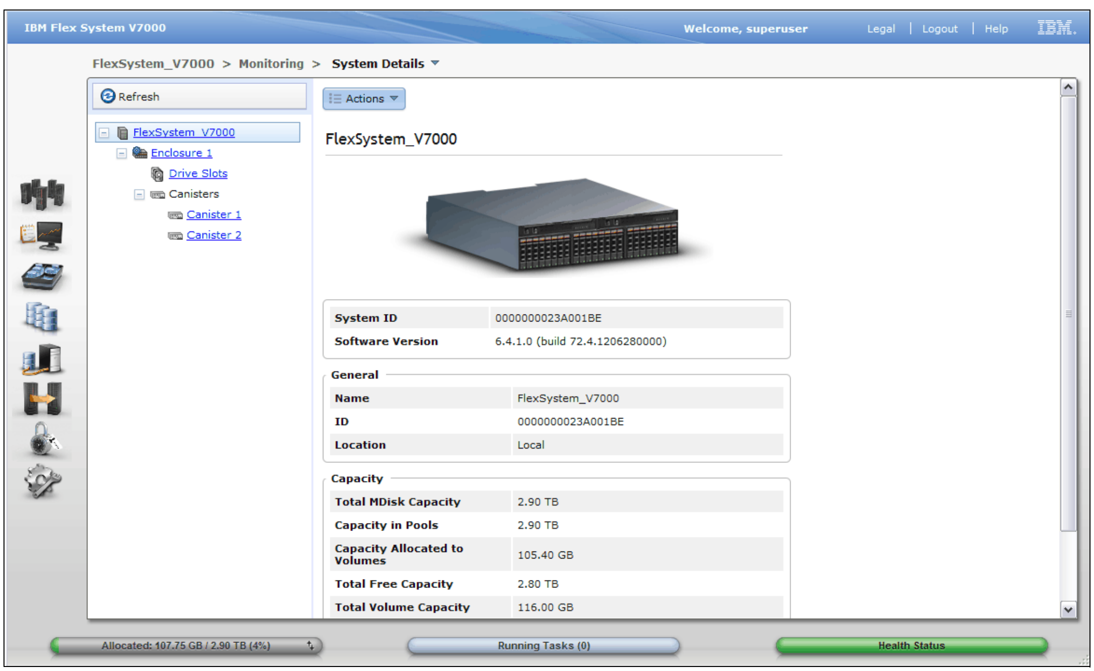


Figure 4-23 System details view in management GUI

13. You may continue to configure additional functions and features for your environment in order to meet your implementation requirements.

4.4 System Management

System management allows for troubleshooting and management tasks using the management graphical user interface (GUI), this includes checking the status of the storage server components, updating the firmware, monitoring events, and managing the storage server. It offers advanced functions, such as FlashCopy, Volume Mirroring, and Remote Mirroring.

A command-line interface (CLI) for the IBM Flex System V7000 Storage Node is available as well.

IBM Flex System V7000 Storage Node is managed through FSM or CMM using the management GUI or a command-line interface (CLI) over an Ethernet connection. For more information, refer to Chapter 3, “Systems Management” on page 81.

Note: Management IP addresses that are assigned to a Flex System V7000 Storage Node are different from iSCSI IP addresses and are used for different purposes. If iSCSI is used, iSCSI addresses are assigned to node ports. On the configuration node, a port has multiple IP addresses active at the same time

4.4.1 Graphical User Interface (GUI)

IBM Flex System V7000 Storage Node graphical user interface (GUI) is designed to simplify storage management and provide a fast and more efficient management tool. A supported web browser should be used for GUI access to configure, manage, and troubleshoot the IBM Flex System V7000 Storage Node. It is also used primarily to configure RAID arrays and logical drives, assign logical drives to hosts, replace and rebuild failed disk drives, and expand the logical drives.

Note: You might need to enable JavaScript in your browser. Additionally, if you are using Firefox, under Advanced JavaScript Settings, you need to click Disable or replace context menus and allow cookies

After the first-time setup procedure is complete, open your web browser on a workstation and point it to your IBM Flex System V7000 Storage Node management IP address (Figure 4-24 on page 168) and complete the steps guided by the Initial Setup wizard. The default login is superuser and the password is passw0rd.

Note: If management IP has been configured for your IBM Flex System V7000 Storage Node, the system name you specified will appear in the login window, as shown in Figure 4-24 on page 168. If a longer system name is used you may need to hover over the system name to view the full system name, this is due to limited display space.

Figure 4-24 IBM Flex System V7000 Storage Node GUI logon screen

Figure 4-25 shows a overview of IBM Flex System V7000 Storage Node GUI.

Figure 4-25 IBM Flex System V7000 Storage Node management GUI Welcome page

For more details on how to use IBM Flex System V7000 Storage Node management GUI, see Chapter 5, “IBM Flex System V7000 Storage Node Graphical user interface overview” on page 175.

4.4.2 Launch IBM Flex System V7000 Storage Node GUI from CMM

When a new control enclosure is detected in the IBM Flex System Enterprise Chassis, the Chassis Management Module (CMM) recognizes the new hardware and starts the initialization process. IBM Flex System V7000 Storage Node Management GUI can be

launched and accessed via CMM after the management IP been assigned to your new Flex System V7000 storage node.

To start, log in to the CMM GUI and navigate to the chassis map. Right Click on the canister you wish to launch and select **Launch Storage Node Controller Console** to access storage node management GUI. See Figure 4-26

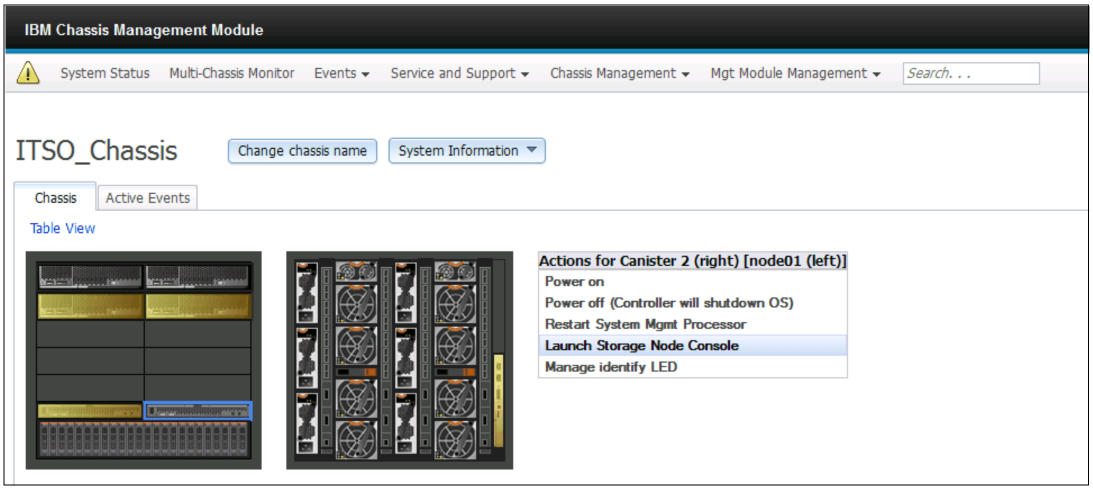


Figure 4-26 To launch storage node controller console via chassis map

To Launch the IBM Flex System V7000 Storage Node management GUI, select the Interface IP address (Figure 4-27) and click **Launch** to start the Storage Node Controller Console.

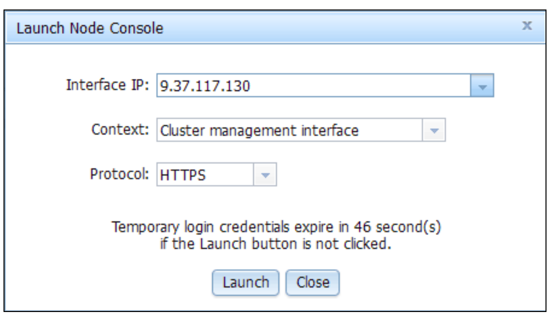


Figure 4-27 Launch to start the Storage Node Controller Console

4.5 Service Assistant

The primary use of the Service Assistant is when a node canister in the IBM Flex System V7000 Storage Node enclosure is in service state. The node canister might be in service state because it has a hardware issue, has corrupted data, or has lost its configuration data. The node canister cannot be active as part of a system while it is in service state.

Attention: Perform service actions on node canisters only when directed to do so by the fix procedures. If used inappropriately, the service actions that are available through the Service Assistant can cause loss of access to data or even data loss.

The storage system management GUI operates only when there is an online system. Use the Service Assistant if you are unable to create a system or if all node canisters in a system are in service state. Use the Service Assistant in the following situations:

- When you cannot access the system from the management GUI and you cannot access the Flex System V7000 Storage Node to run the recommended actions
- When the recommended action directs you to use the Service Assistant.

> **Note:** The Service Assistant does not provide any facilities to help you service expansion enclosures. Always service the expansion enclosures by using the management GUI.

You can also perform the following service-related actions using Service Assistant:

- Collect logs to create and download a package of files to send to support personnel.
- Remove the data for the system from a node.
- Recover a system if it fails.
- Install a software package from the support site or rescue the software from another node.
- Upgrade software on node canisters manually versus performing a standard upgrade procedure.
- Configure a control enclosure chassis after replacement.
- Change the Service IP address that is assigned to Ethernet port 1 for the current node canister.
- Install a temporary SSH key if a key is not installed and CLI access is required.
- Restart the services used by the system.

4.5.1 Change Service IP address

There are multiple methods that you can use to change the Service IP address of a storage node canister and which method to use depends on the status of the system and the other node canisters in the system. Follow the methods in the order shown until you are successful in setting the IP address to the required value.

> **Note:** You can set an IPv4 address, an IPv6 address, or both, as the service address of a node. Enter the required address correctly. If you set the address to 0.0.0.0 or 0000:0000:0000:0000:0000:0000:0000, you disable the access to the port on that protocol.

Change Service IP via management GUI

Use the control enclosure management GUI when the system is operating and the system is able to connect to the node canister with the Service IP address that you want to change, as shown in Figure 4-28 on page 171.

1. Select **Settings** → **Network** from the navigation.
2. Be sure to select the correct node to configure.
3. Select Specify IP address and change it to the new IP
4. Click Save to complete the panel.

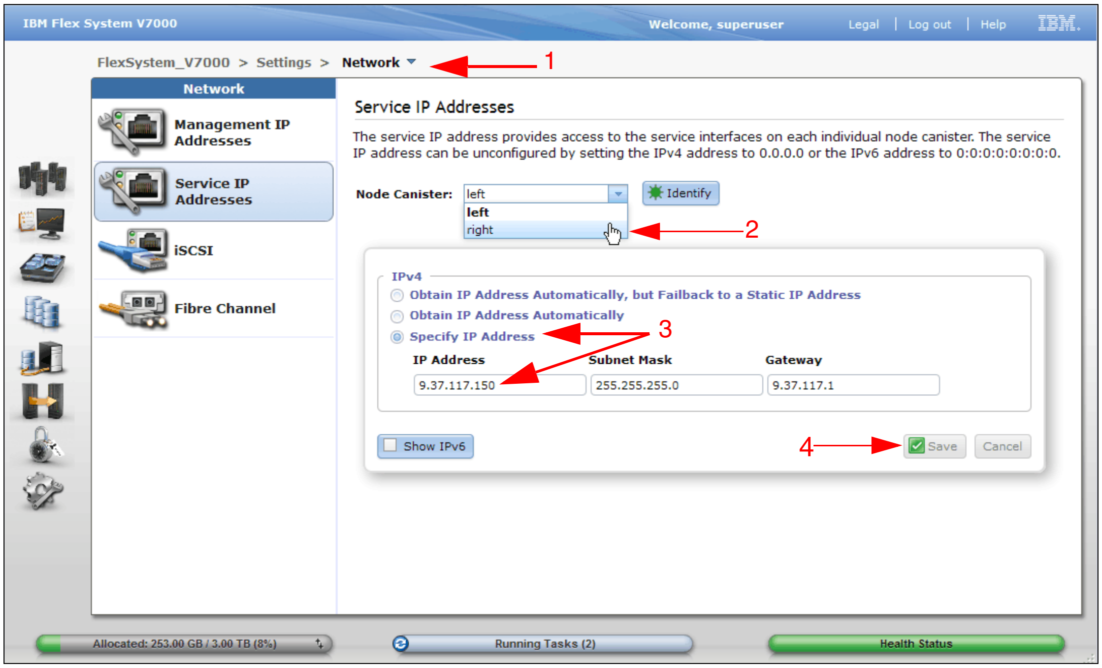


Figure 4-28 Change Service IP via management GUI procedure

Change Service IP via CMM

If CMM in the chassis is operating and reachable, you can use the CMM to change the Service IP address of a node canister, as shown in Figure 4-29 on page 172.

> **Note:** If, for example, the canister has failed to obtain a DHCP address and a default Service IP address in the range between 192.168.70.131 and 192.168.70.144 is displayed, the canister has failed to obtain a DHCP address.

1. Open the CMM and navigate to the **Chassis Management** → **Component IP Configuration** page
2. In the list of Storage Nodes, click on the name of the node canister.
3. In the IP address Configuration panel, click the IPv4 tab.
4. Select **Node_Service_If** as the Network Interface and select the Configuration Method you wish to use. If the selected Configuration Method is to Use Static IP address (as shown in Figure 4-29 on page 172), type the Static IP address, Subnet Mask, and Gateway address for the New Static IP Configuration.
5. Click Apply and then Close

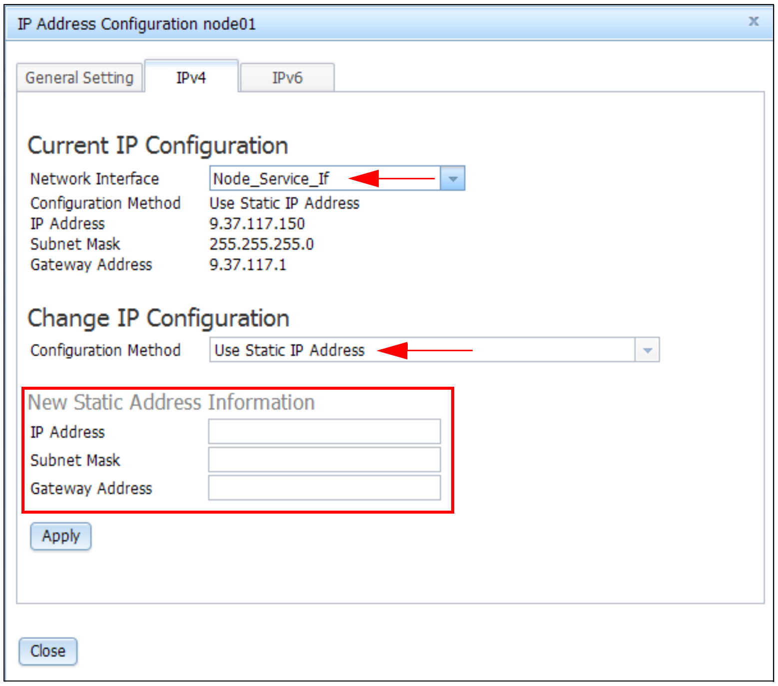


Figure 4-29 Change Service IP via CMM procedure

Change service IP via CLI

These addresses are not set during the installation of a Flex System V7000 Storage Node system, but you can set these IP addresses later by using the **chserviceip** CLI command.

4.6 Command-Line interface (CLI)

The Flex System V7000 Storage Node command-line interface (CLI) is a collection of commands that you can use to manage a Flex System V7000 Storage Node.

You can use the CLI to perform the following functions:

- Setup of the clustered system, its nodes, and the I/O groups
- Analyze error logs (event logs)
- Setup and maintenance of managed disks (MDisk) and storage pools
- Setup and maintenance of client public SSH keys on the clustered system
- Setup and maintenance of volumes
- Setup of logical host objects
- Map volumes to hosts
- Navigate from managed hosts to volumes and to MDisks, and the reverse direction up the chain
- Set up and start Copy Services:,
  - FlashCopy and FlashCopy consistency groups
  - Synchronous Metro Mirror and Metro Mirror consistency groups
  - Asynchronous Global Mirror and Global Mirror consistency groups

The CLI commands use the Secure Shell (SSH) connection between the SSH client software on the host system and the SSH server on the Flex System V7000 Storage Node. Install and set up SSH client software on each system that you plan to use to access the CLI. Authenticate to the system using a password. If you require command line access without

entering a password, use an SSH public key. Then store the SSH public key for each SSH client on the Flex System V7000 Storage Node.

> **Note:** After the first SSH public key is stored, you can add additional SSH public keys using either the management GUI or the CLI.

You can connect to the Flex System V7000 Storage Node CLI using PuTTY (Figure 4-30).

```
login as: superuser
superuser@9.37.117.130's password:
Last login: Mon Jul 23 21:08:45 2012 from 9.80.11.49
IBM_4939:FlexSystem_V7000:superuser>sainfo lsservicenodes
panel_name cluster_id       cluster_name     node_id node_name relation node_sta
tus error_data
01-1       0000000023A001BE FlexSystem_V7000 1       node1     local    Active
01-2       0000000023A001BE FlexSystem_V7000 2       node2     partner  Active
IBM_4939:FlexSystem_V7000:superuser>
```

Figure 4-30 Connect to the CLI using PuTTY

Detailed Command Line Interface setup procedure for the IBM Flex System V7000 Storage Node is available in Appendix A, "CLI Setup and Configuration" on page 573 and can also be found at the following address:

http://pic.dhe.ibm.com/infocenter/storwize/ic/index.jsp?topic=/com.ibm.storwize.v7000.doc/svc_clicommandscontainer_229g0r.html

4.6.1 Record system access information

It is important that anyone who has responsibility for managing the system know how to connect to and log on to the system. Give attention to those times when the normal system administrators are not available because of vacation or illness. Record the following information and ensure that authorized people know how to access the system information as shown in Table 4-1:

Table 4-1 Record log on information

Item	Value	Notes
The management IP address for the GUI and CLI		
The management user ID (the default is admin)		
The management user ID password (the default is admin)		
The control enclosure management IP address		
Control enclosure Service IP address: node canister 1		
Control enclosure Service IP address: node canister 2		

Item	Value	Notes
The control enclosure superuser password (the default is passw0rd)		

5

IBM Flex System V7000 Storage Node Graphical user interface overview

This chapter provides an overview of the graphical user interface of the IBM Flex System V7000 Storage Node and shows how to use the navigation tools.

5.1 Overview of IBM Flex System V7000 Storage Node management software

The IBM Flex System V7000 Storage Node can be managed from either the IBM Flex System Manager Node (FSM) or from the built in Graphical User Interface (GUI) which is a web browser based management tool.

The web based Graphical User Interface is designed to simplify storage management and provide a fast and more efficient management tool. It is loosely based on the IBM System Storage XIV software and has a similar look and feel.

JavaScript: You may need to enable JavaScript in your browser. Additionally, if you are using Firefox, under Advanced JavaScript Settings, you need to click **Disable or replace context menus** and allow cookies.

For more information of managing the IBM Flex System V7000 Storage Node through FSM, refer to Chapter 3, "Systems Management" on page 81.

5.1.1 Access to the Graphical User Interface

To log on to the Graphical User Interface, point your web browser at the IP address that was set during the initial setup of the IBM Flex System V7000 Storage Node. The login window opens (Figure 5-1).

Figure 5-1 Login window

After you have logged in successfully, the **Home** window opens showing the **Overview** screen (Figure 5-2).

Figure 5-2 Home menu showing the Overview window

5.1.2 Graphical User Interface layout

This Graphical User Interface (GUI) has three main sections for navigating through the management tool. On the far left of the window are eight *function icons*. The eight function icons represent:

- The Home menu
- The Monitoring menu
- The Pools menu
- The Volumes menu
- The Hosts menu
- The Copy Services menu
- The Access menu
- The Settings menu

In the middle of the window is a diagram illustrating the existing configuration. Clicking the function icons in this area provides extended help references, including a link to a short video presentation to explain the topic in more detail. This video is not a navigation tool, but rather an extended help window that includes configuration information.

At the bottom of the window are three status indicators. Clicking any of them provides more detailed information about the existing configuration of the IBM Flex System V7000 Storage Node solution. Click any of these function icons to expand them and minimize them as required.

Figure 5-3 shows the main areas of the GUI Home Overview.

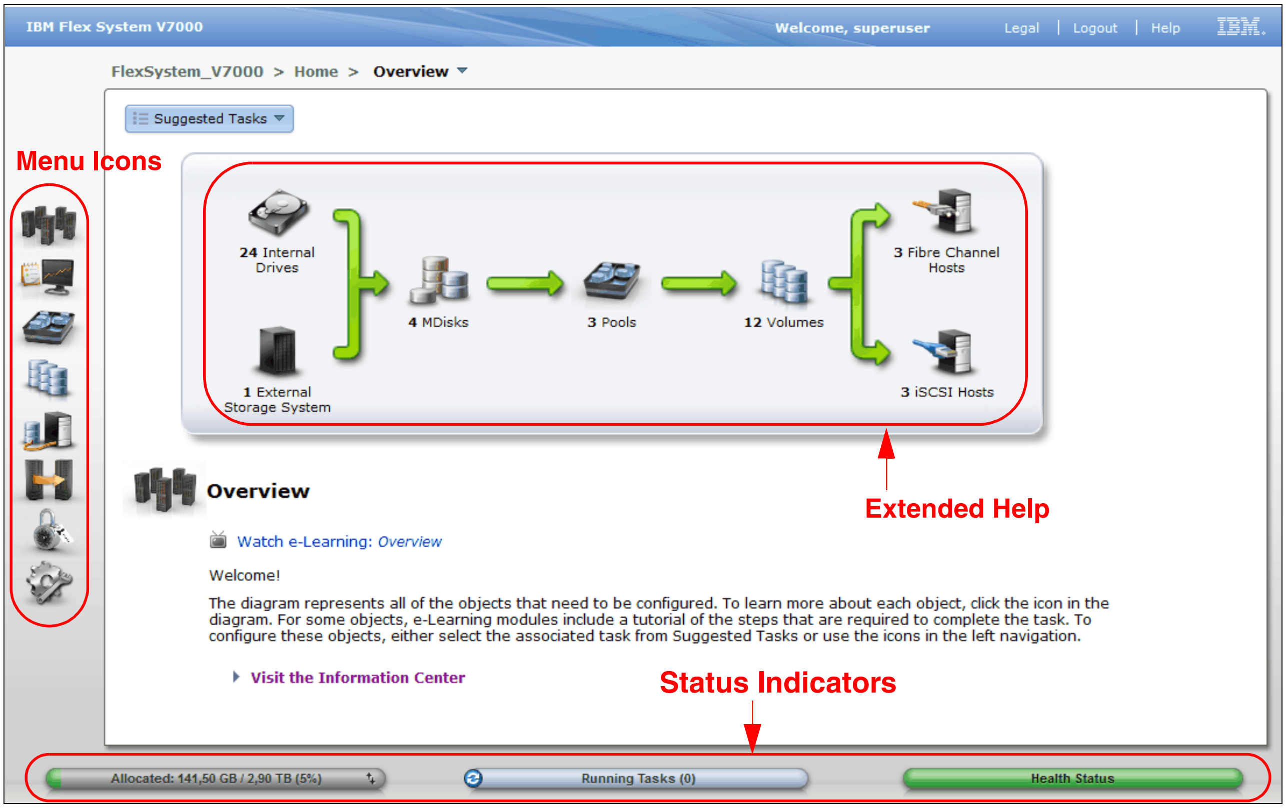


Figure 5-3 Main areas

5.1.3 Navigation

Navigating around the management tool is simple. You can hover the cursor over one of the eight function icons on the left side of the window, which highlights the function icon and shows a list of options. You can then move the cursor to the wanted option and click it (Figure 5-4 on page 179).

Figure 5-4 Navigation using the menu options

Figure 5-5 shows a list of the IBM Flex System V7000 Storage Node Software function icons and the associated menu options.

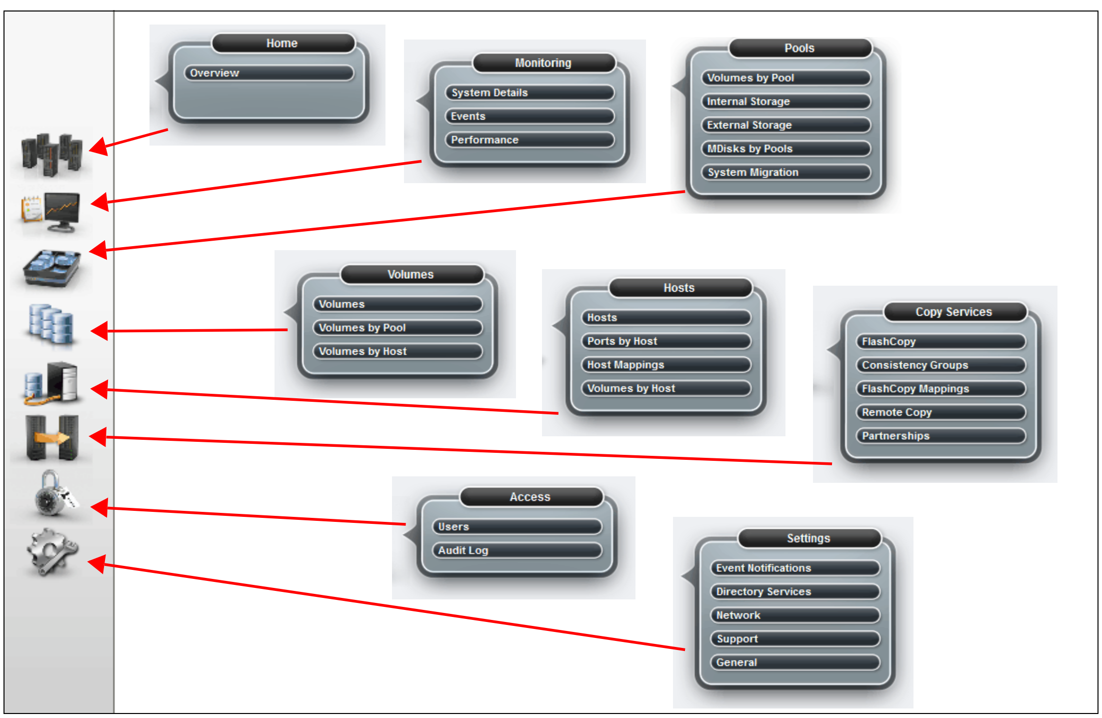


Figure 5-5 IBM Flex System V7000 Storage Node Software - Menu options

If you for example click the **Volumes** menu, you can change the window's view (Figure 5-6). This action also applies to any other menu options.

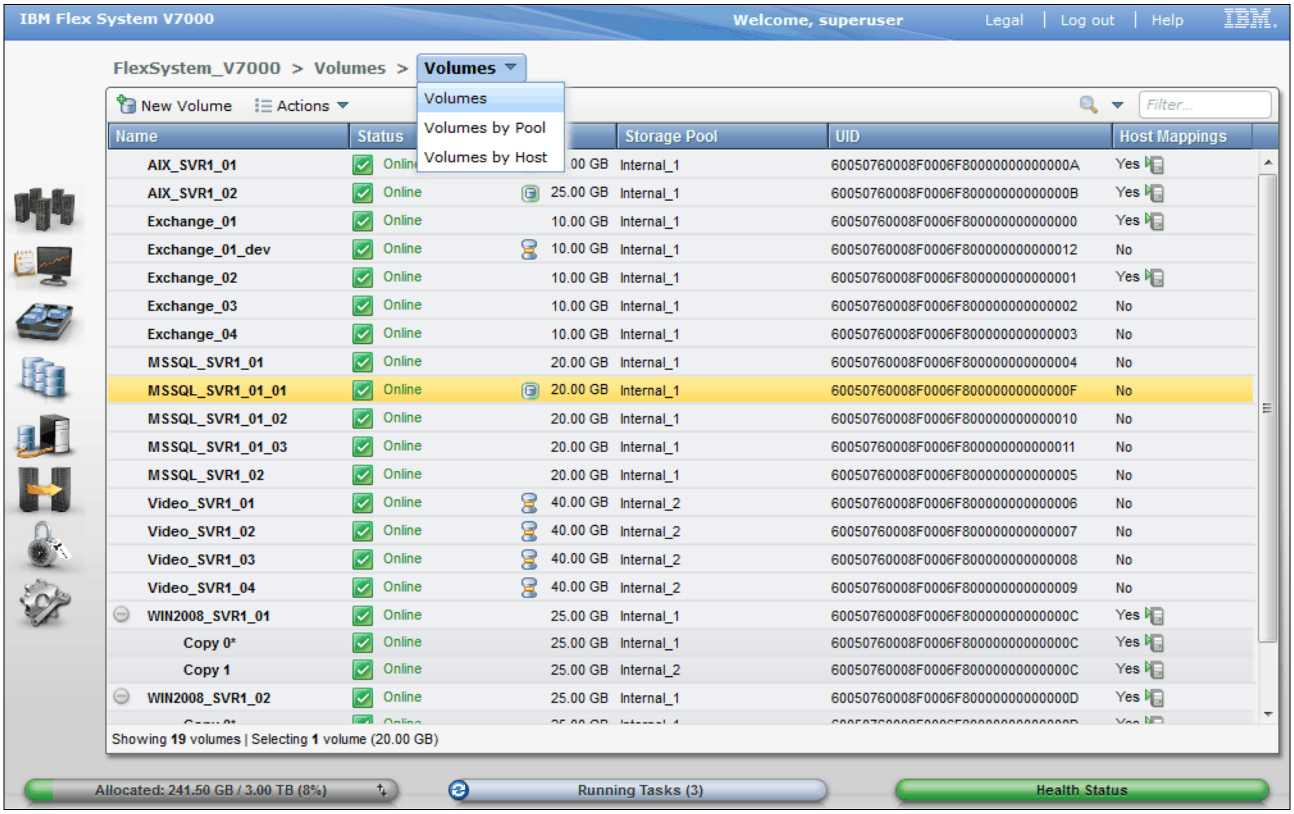


Figure 5-6 Navigation using the change view

5.1.4 Multiple selections

The new management tool also lets you select multiple items by using a combination of the Shift or Ctrl keys. To select multiple items in a display, click the first item, press and hold the Shift key, and click the last item in the list you require. All the items in between those two items are selected (Figure 5-7 on page 181).

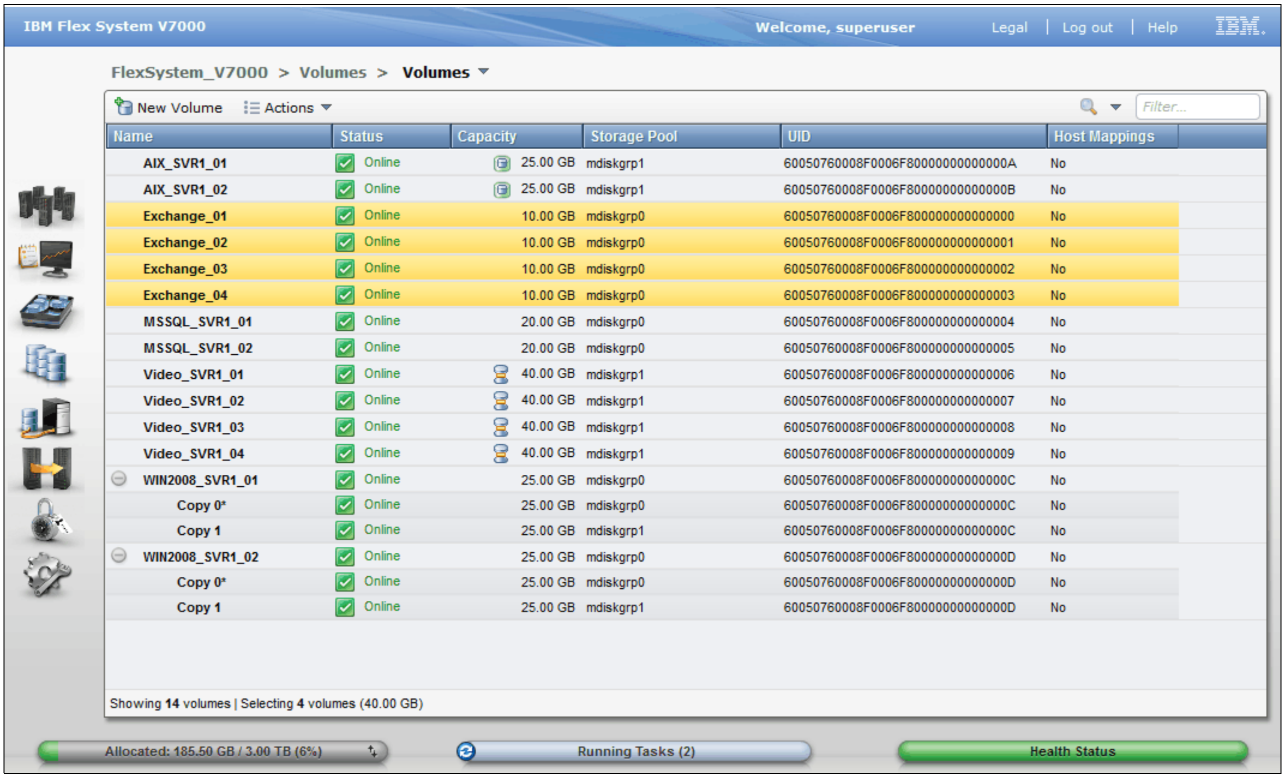
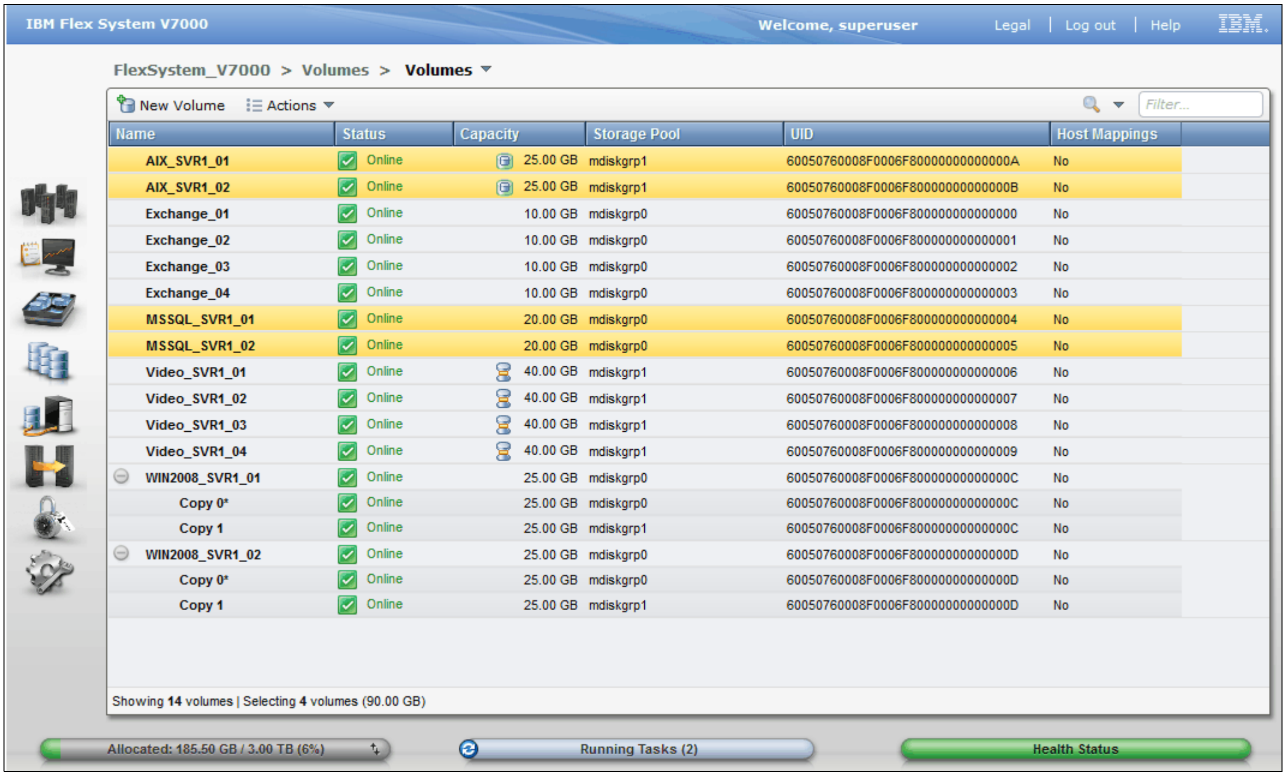


Figure 5-7 Multiple selections using the Shift key

If you want to select multiple items that are not in sequential order, click the first item, press and hold the Ctrl key, and click the other items you require (Figure 5-8).



Name	Status	Capacity	Storage Pool	UID	Host Mappings
AIX_SVR1_01	Online	25.00 GB	mdiskgrp1	60050760008F0006F8000000000000A	No
AIX_SVR1_02	Online	25.00 GB	mdiskgrp1	60050760008F0006F8000000000000B	No
Exchange_01	Online	10.00 GB	mdiskgrp0	60050760008F0006F80000000000000	No
Exchange_02	Online	10.00 GB	mdiskgrp0	60050760008F0006F80000000000001	No
Exchange_03	Online	10.00 GB	mdiskgrp0	60050760008F0006F80000000000002	No
Exchange_04	Online	10.00 GB	mdiskgrp0	60050760008F0006F80000000000003	No
MSSQL_SVR1_01	Online	20.00 GB	mdiskgrp0	60050760008F0006F80000000000004	No
MSSQL_SVR1_02	Online	20.00 GB	mdiskgrp0	60050760008F0006F80000000000005	No
Video_SVR1_01	Online	40.00 GB	mdiskgrp1	60050760008F0006F80000000000006	No
Video_SVR1_02	Online	40.00 GB	mdiskgrp1	60050760008F0006F80000000000007	No
Video_SVR1_03	Online	40.00 GB	mdiskgrp1	60050760008F0006F80000000000008	No
Video_SVR1_04	Online	40.00 GB	mdiskgrp1	60050760008F0006F80000000000009	No
WIN2008_SVR1_01	Online	25.00 GB	mdiskgrp0	60050760008F0006F8000000000000C	No
Copy 0*	Online	25.00 GB	mdiskgrp0	60050760008F0006F8000000000000C	No
Copy 1	Online	25.00 GB	mdiskgrp1	60050760008F0006F8000000000000C	No
WIN2008_SVR1_02	Online	25.00 GB	mdiskgrp0	60050760008F0006F8000000000000D	No
Copy 0*	Online	25.00 GB	mdiskgrp0	60050760008F0006F8000000000000D	No
Copy 1	Online	25.00 GB	mdiskgrp1	60050760008F0006F8000000000000D	No

Figure 5-8 Multiple selections using the Ctrl key

5.1.5 Status Indicators menus

Other useful tools are the Status Indicators that appear at the bottom of the window. These indicators provide information about Capacity usage, Compression ratio, Running Tasks and the Health Status of the system. The Status Indicators are visible from all screens in the IBM Flex System V7000 Storage Node GUI.

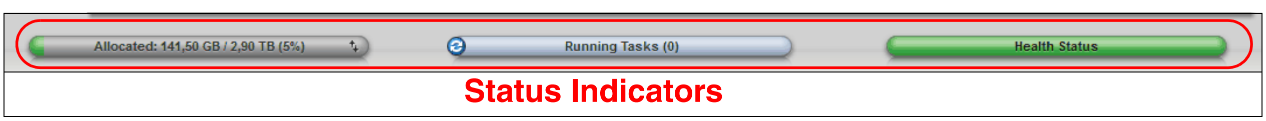


Figure 5-9 Show running task details window

For example you can view the details of the running tasks by clicking the button on the status indicator bars at the bottom of the window (Figure 5-10 on page 183).

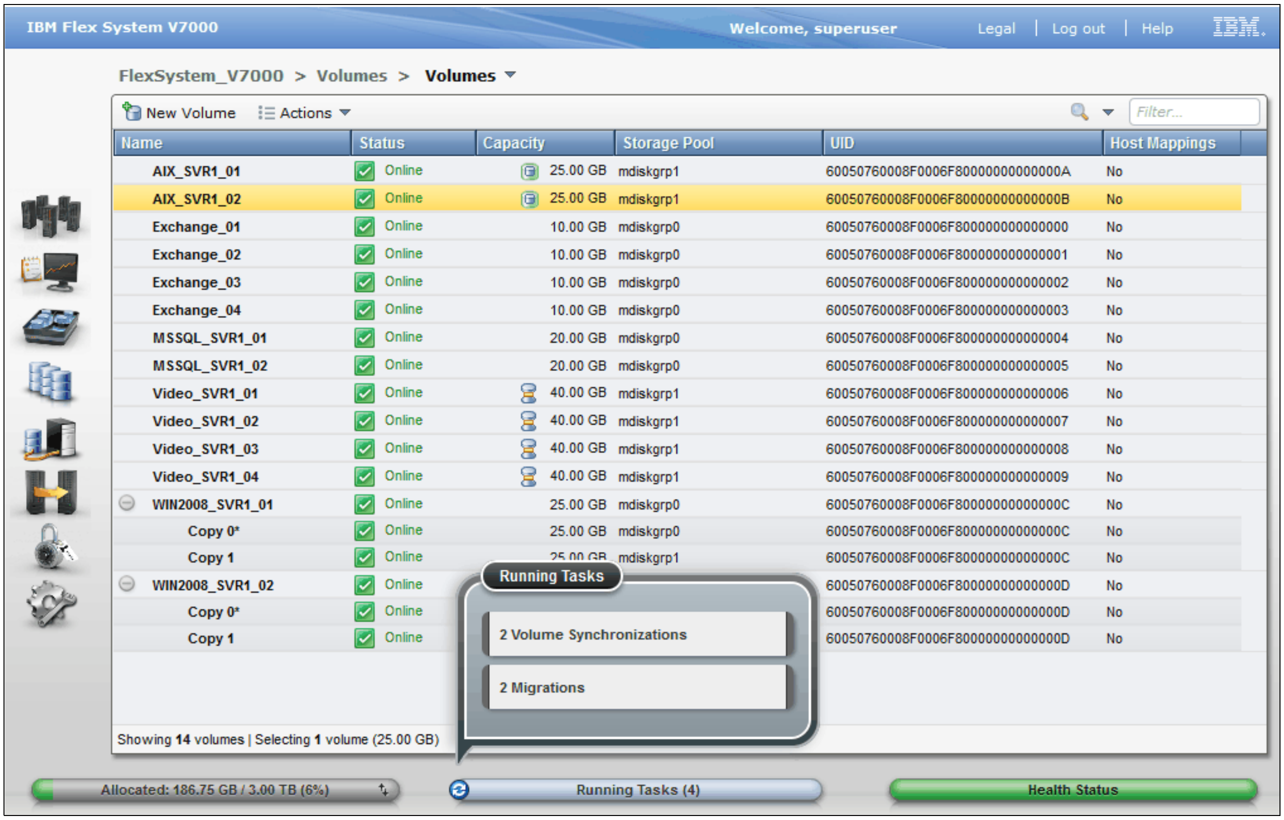


Figure 5-10 Status Indicators

Figure 5-11 shows additional information about Recently Completed tasks when it opens from the Running Tasks menu.

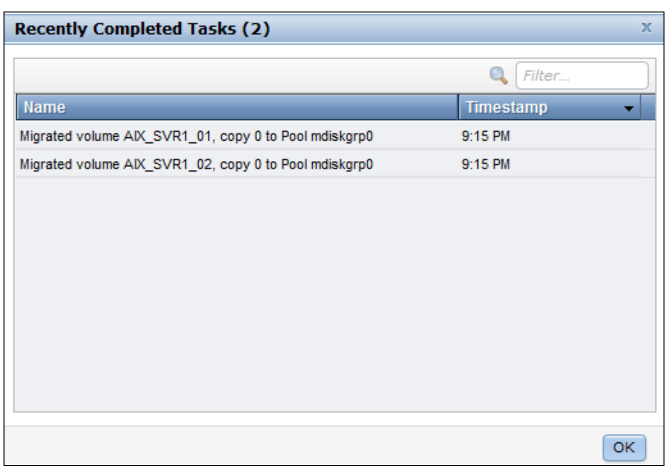


Figure 5-11 Status Indicators - Recently Completed tasks

5.2 Home menu

The IBM Flex System V7000 Storage Node management software provides an efficient and quick mechanism for navigating between the various different functions. Clicking one of the eight function icons on the left side of the window causes a menu option to open that allows you to navigate directly to the selected option (Figure 5-12).

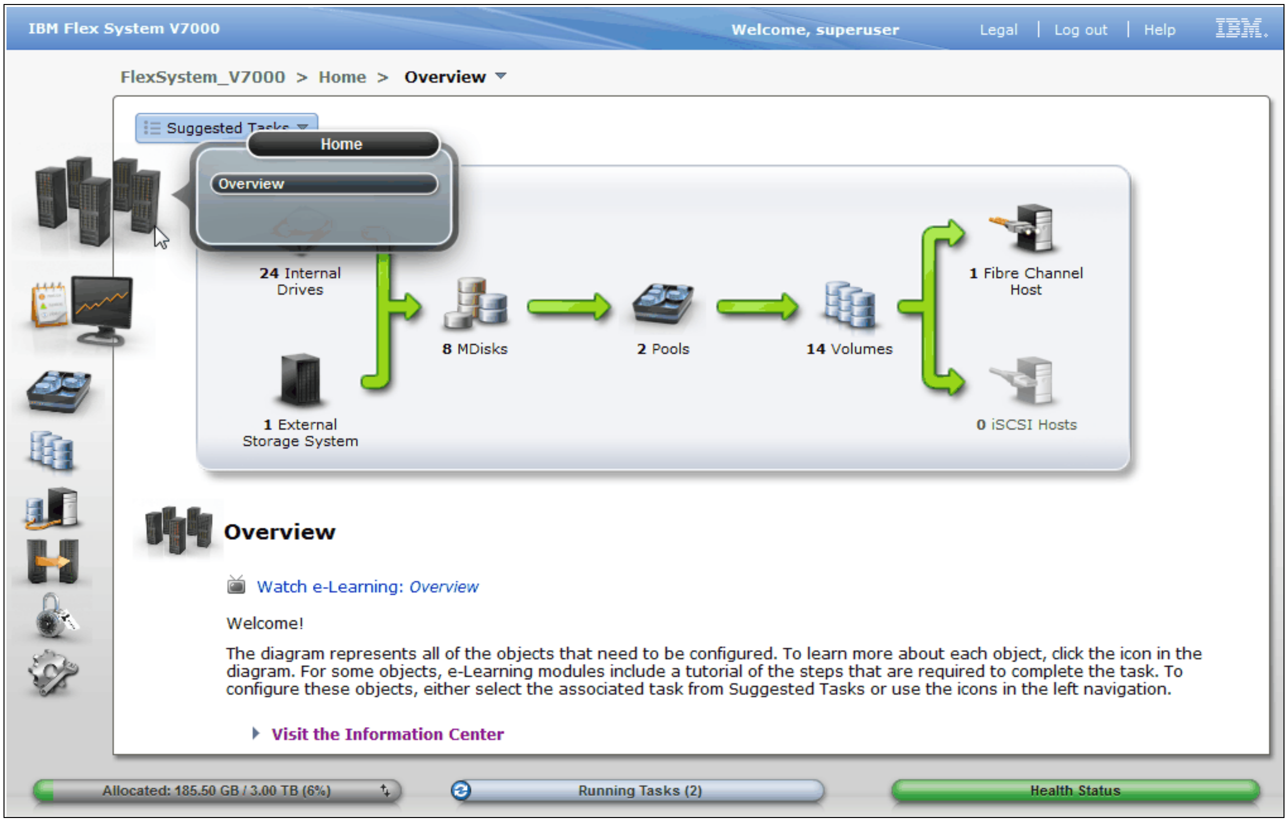


Figure 5-12 Navigate to the Home menu

5.3 Monitoring menu

In this section, we describe the Monitoring menu and its options.

5.3.1 Monitoring System Details menu

Figure 5-13 on page 185 shows the Monitoring → System Details menu in where detailed information about firmware, World Wide Names, Capacity usage and other important information can be found.

Figure 5-13 Monitoring System Details menu

After clicking on the System Details, window shown in Figure 5-14 on page 186 is displayed. This shows the upper level of the Details section of the IBM Flex System V7000 Storage Node. From here the firmware level, host name as well as capacity usage can be reviewed and changed. Pressing the **Actions** button here gives a number of options e.g. the system can be Identified with a beacon on the front of the system and system firmware can be upgraded.

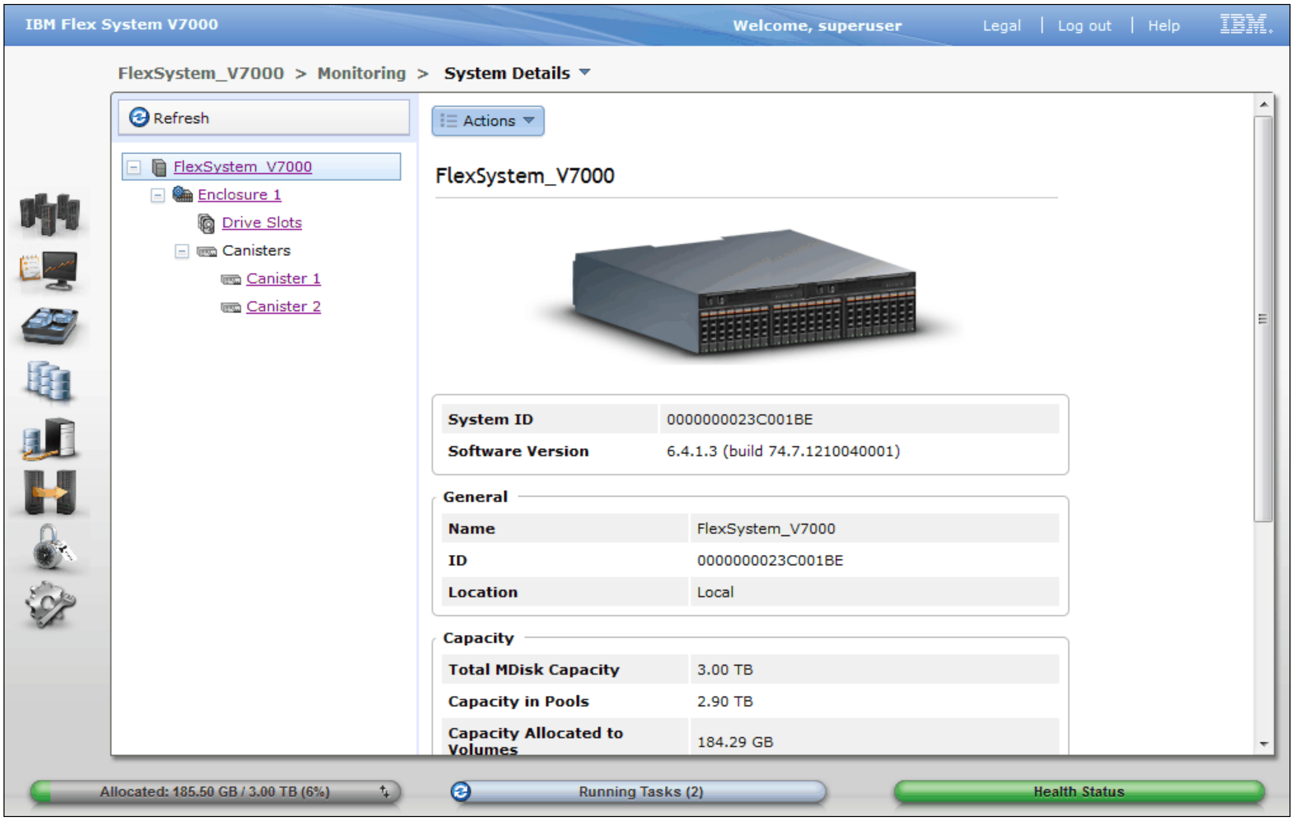


Figure 5-14 System Details

The IBM Flex System V7000 Storage Node Control Enclosure contains Array controllers called Canisters as well as up to 24 disk drives. Figure 5-15 on page 187 shows properties for the Control Enclosure including the system Serial number and Machine Type and Model number.

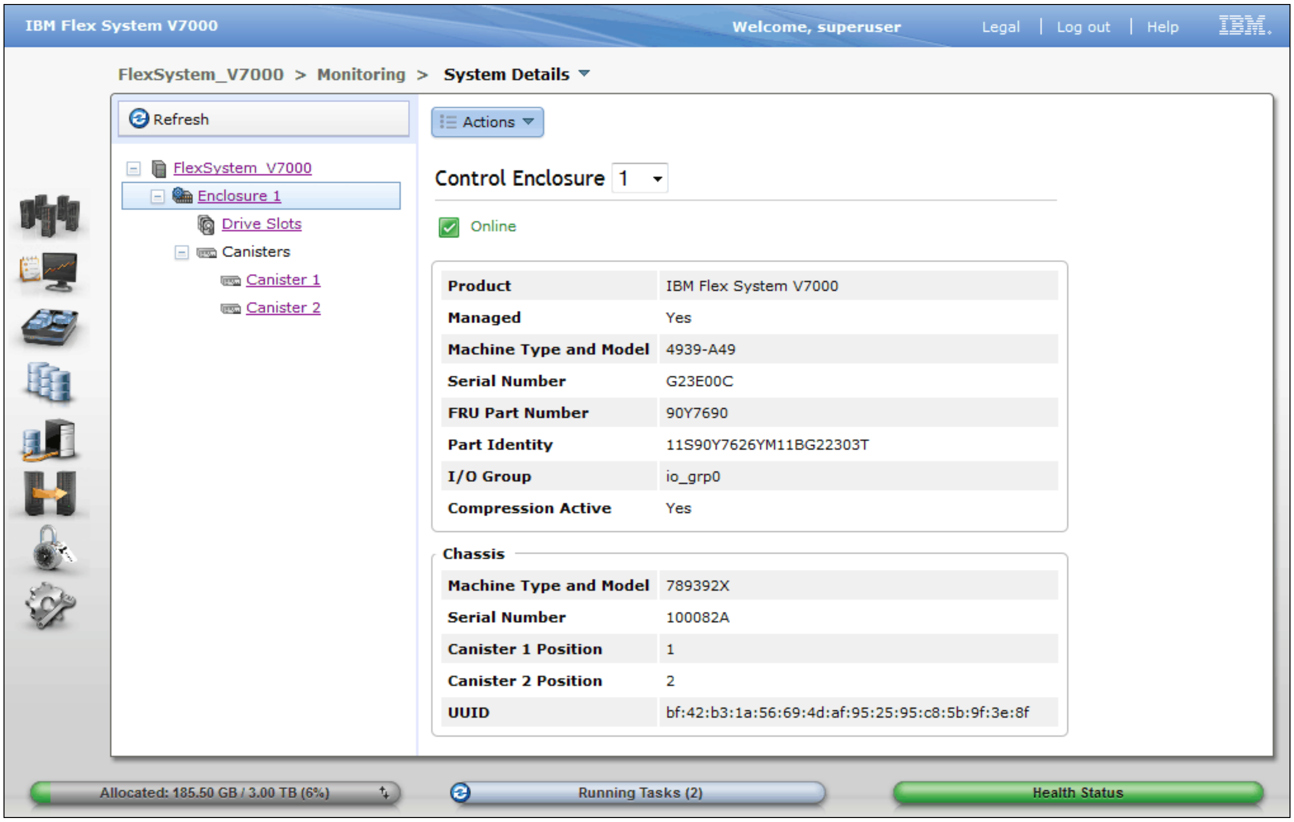


Figure 5-15 System Details - Enclosure

Figure 5-16 on page 188 shows the top section of the properties for the left Canister including Node name, SAS-port status and the iSCSI Initiator name (IQN Name).

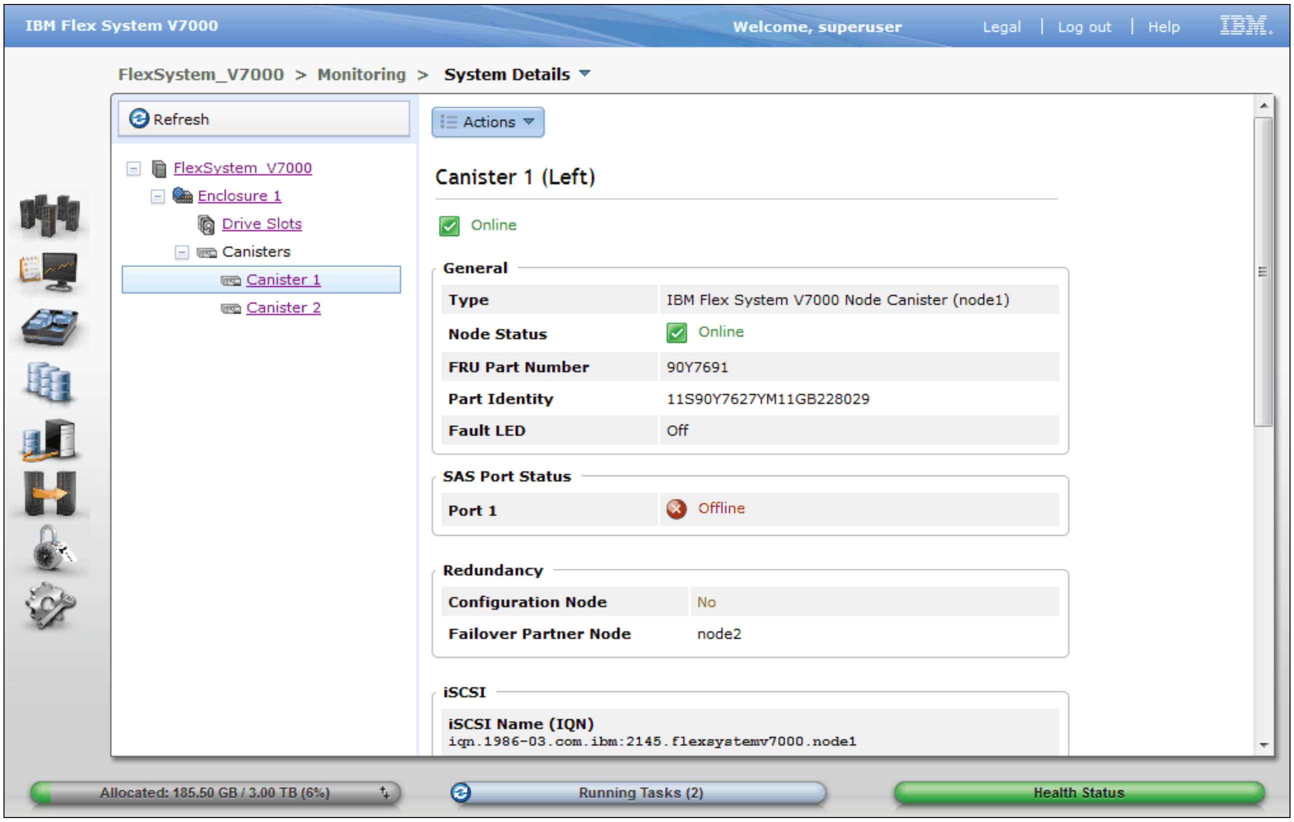


Figure 5-16 System Details - Canisters

In the Figure 5-16, the SAS Port 1 is Offline as the control enclosure is not currently connected to any disk enclosures.

Figure 5-17 on page 189 shows the bottom section of the properties for the left Canister where World Wide Names (WWN) can be reviewed. The WWNs are used when the IBM Flex System V7000 Storage Node connects to an external Fibre Channel attached storage system through SAN switches. For these SAN switches appropriate SAN switch zoning must be applied for the external system to communicate with the IBM Flex System V7000 Storage Node.

For more information about SAN zoning for the IBM Flex System V7000 Storage Node check Chapter 11, “SAN Connections and Configuration” on page 431.

For more information about SAN zoning with IBM/Brocade switches refer to *Implementing an IBM b-type SAN with 8 Gbps Directors and Switches*, SG24-6116.

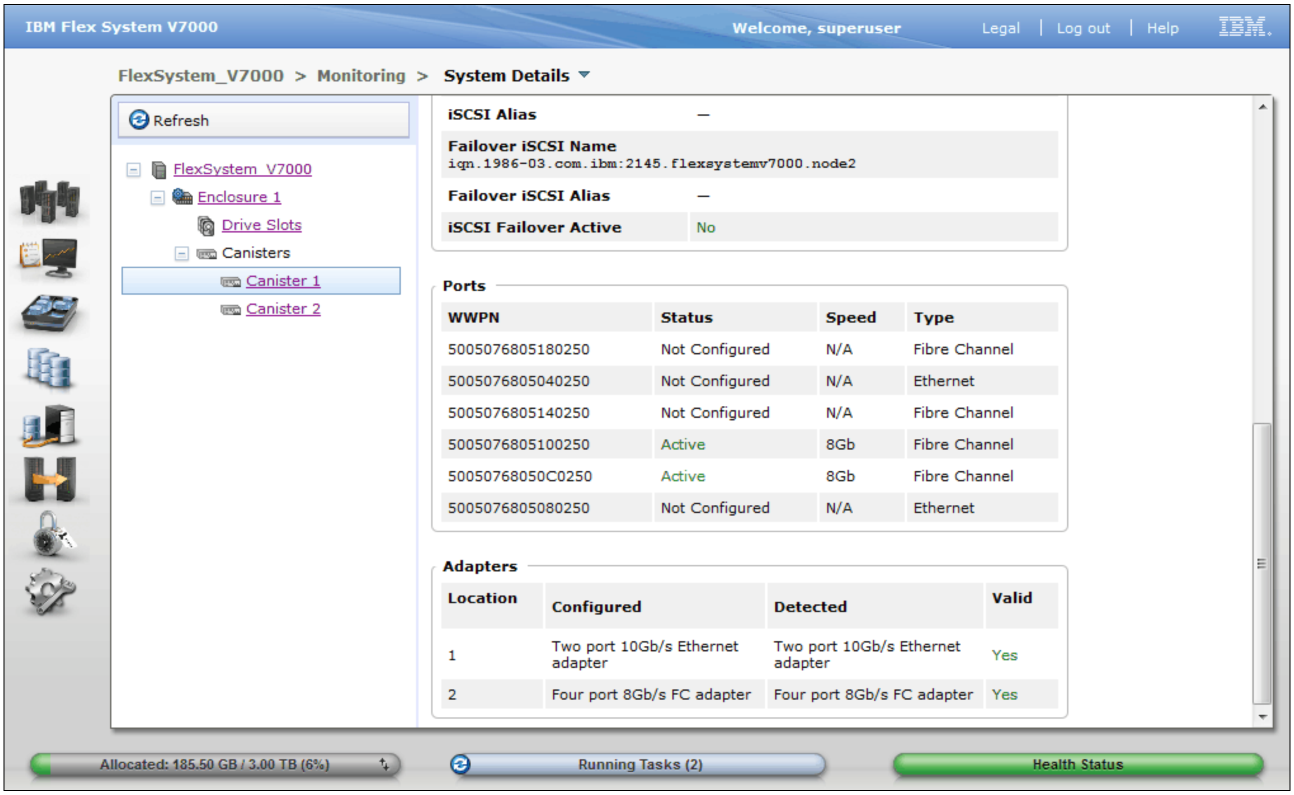


Figure 5-17 System Details - Canisters IO modules and WWNs

Another detail of this window is the Adapters section. This contains information of which Host Interface Cards (HICs) are installed in the system. For more information of Host Interface Cards and in which slots they can be installed refer to Chapter 2, “Introduction to IBM Flex System V7000 Storage Node” on page 39.

5.3.2 Monitoring Events menu

The IBM Flex System V7000 Storage Node may from time to time show “Health Status” status bar as Yellow (Degraded) or even Red (Critical). Events on the IBM Flex System V7000 Storage Node system are logged in the Event Log of the Monitoring Events menu.

An example of one of the most common errors that can be reviewed and fixed from the Event Log is the error 1630 “Number of Device Logins Reduced”. This error happens when an underlying disk array looses connectivity through one or more of its ports that are zoned to and connected to the IBM Flex System V7000 Storage Node. Few possible reasons for that could be a faulty Fibre Channel connection, a SAN-switch that was rebooted or if an underlying disk array controller was taken out for service.

Such events are logged as errors and has to be fixed before theIBM Flex System V7000 Storage Node recovers out of the error condition it was in due to the event.

To navigate to the Event Log hover the cursor over the Monitoring icon and the click on **Events** (Figure 5-18 on page 190).

Figure 5-18 Navigate to the Monitoring Events menu

Recommended actions

In the Events window, messages and errors are displayed. In case any unresolved issues exist the **Next Recommended Actions** section displays the recommended actions and it is possible to run a Fix Procedure. By doing that the IBM Flex System V7000 Storage Node will check if the problem still exists and it will fix the issue if possible. The fix procedure may bring the system out of a Degraded state and into a Healthy state.

Figure 5-19 on page 191 shows that an error exist which has to be fixed. To fix the event, click **Run Fix Procedure**, and a Directed Maintenance Procedure (DMP) starts that walks you through the process to fix that particular event.

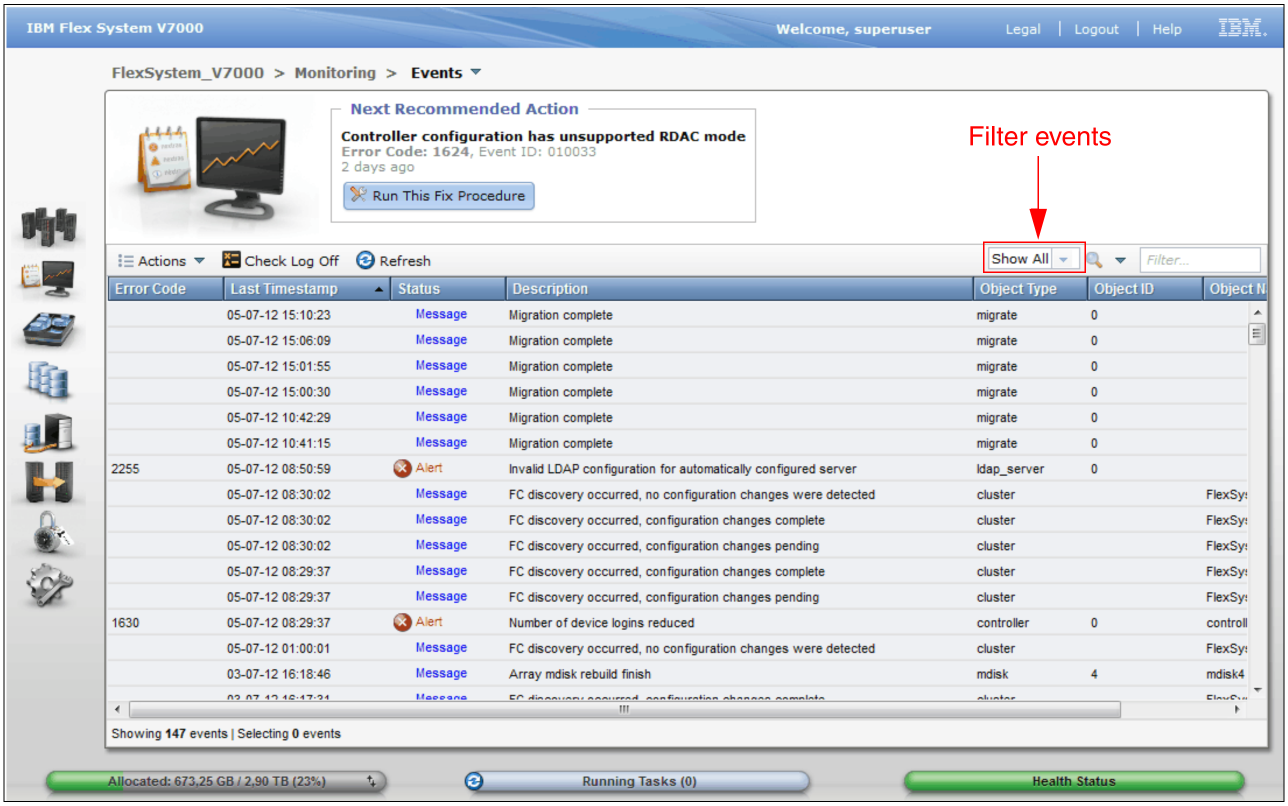


Figure 5-19 Recommended Actions menu

In Figure 5-19, you can also use the option to filter the events and display only the events you need to see.

Another way to fix an event is to right-click the particular event and click **Run Fix Procedure** as shown in Figure 5-20.

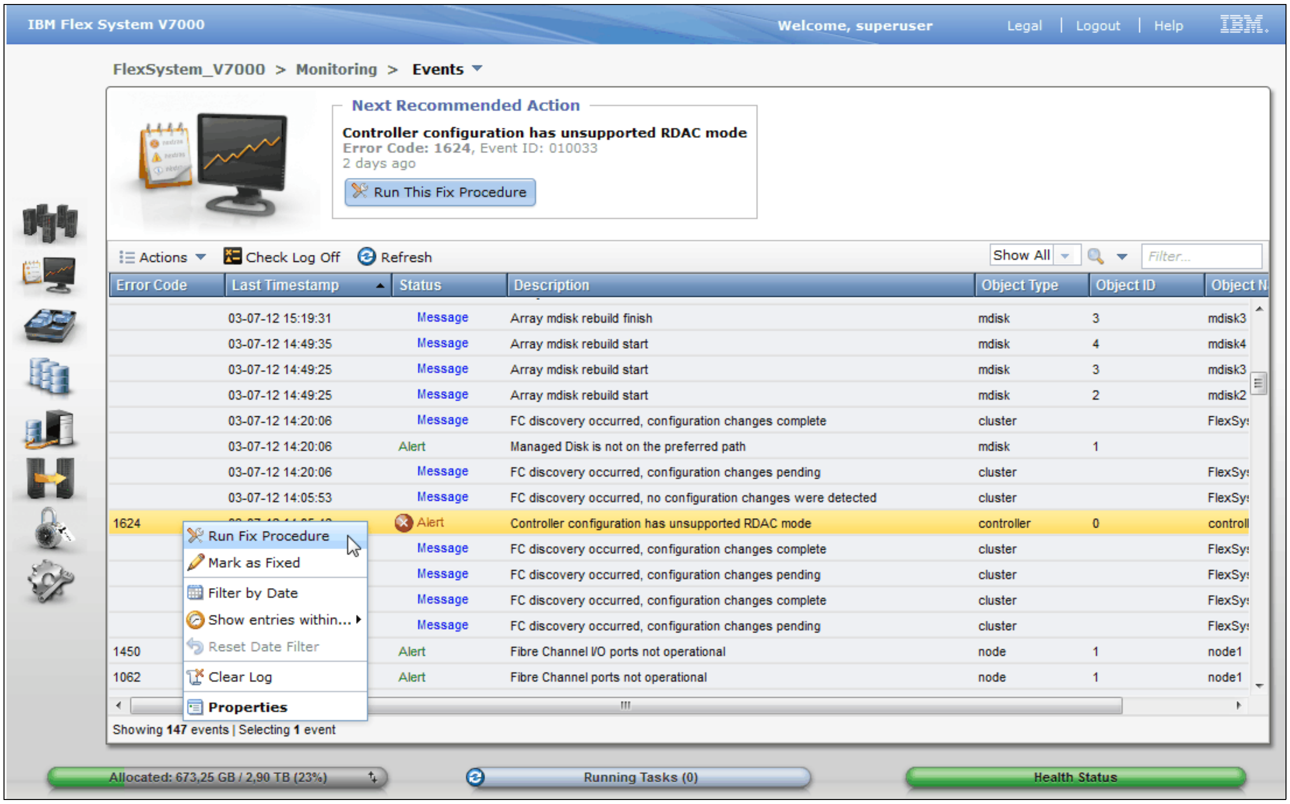


Figure 5-20 Event Fix procedure using right click

To view the details of a specific event, click **Properties** (as shown in Figure 5-20). The Properties window opens as shown in Figure 5-21 on page 193 and shows the details of the error.

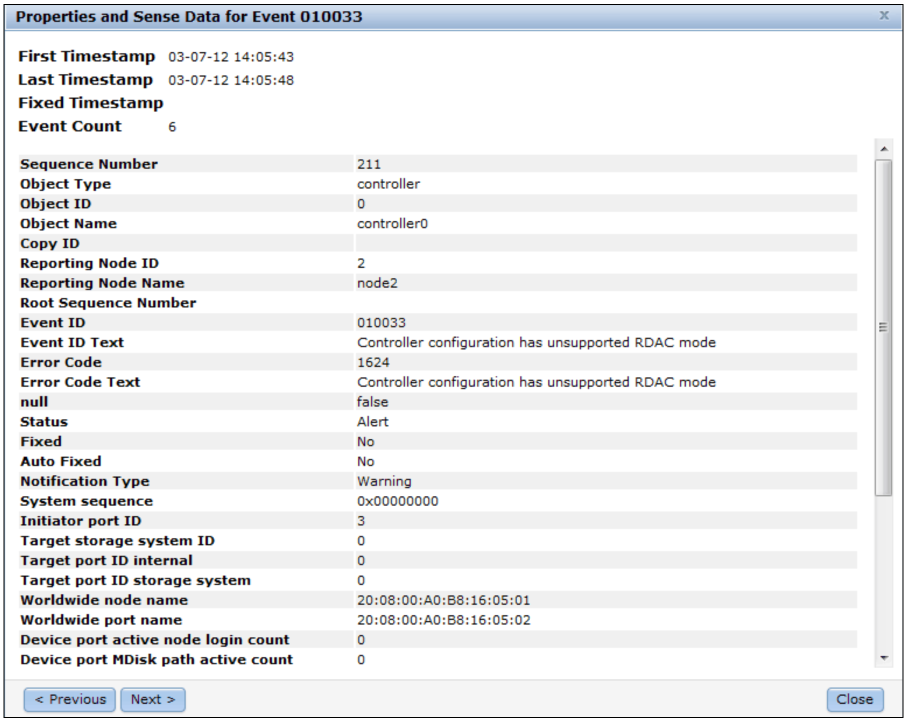


Figure 5-21 Properties of the event

Performing the Fix Procedure

In Figure 5-22 on page 194, we demonstrate how an error is fixed. Before starting the fix procedure we filter events to **Recommended Actions** so that only errors are displayed that require attention. Click **Run This Fix Procedure.**

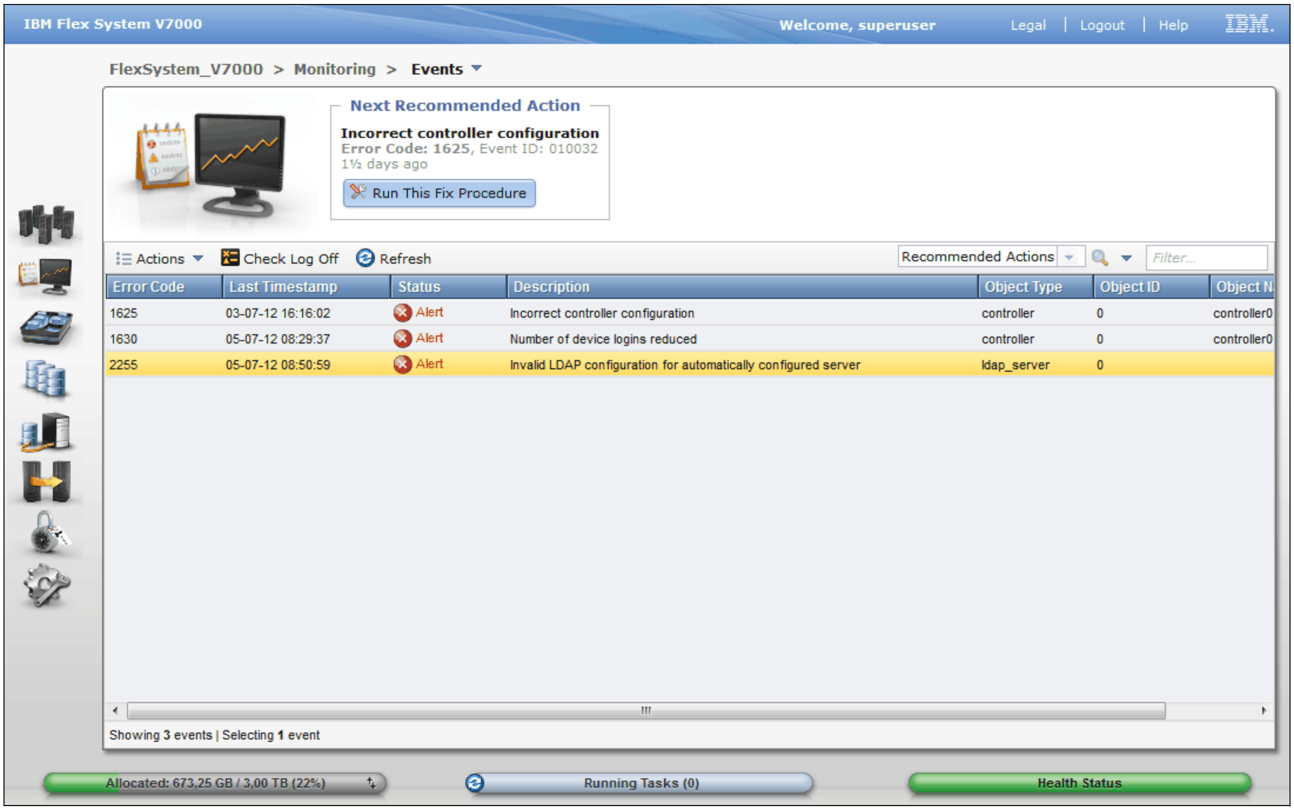


Figure 5-22 Error code 1625 Fix Procedure before fix

Figure 5-23 shows the first step of the fix procedure for the error with Event Code1625 Incorrect Controller Configuration.

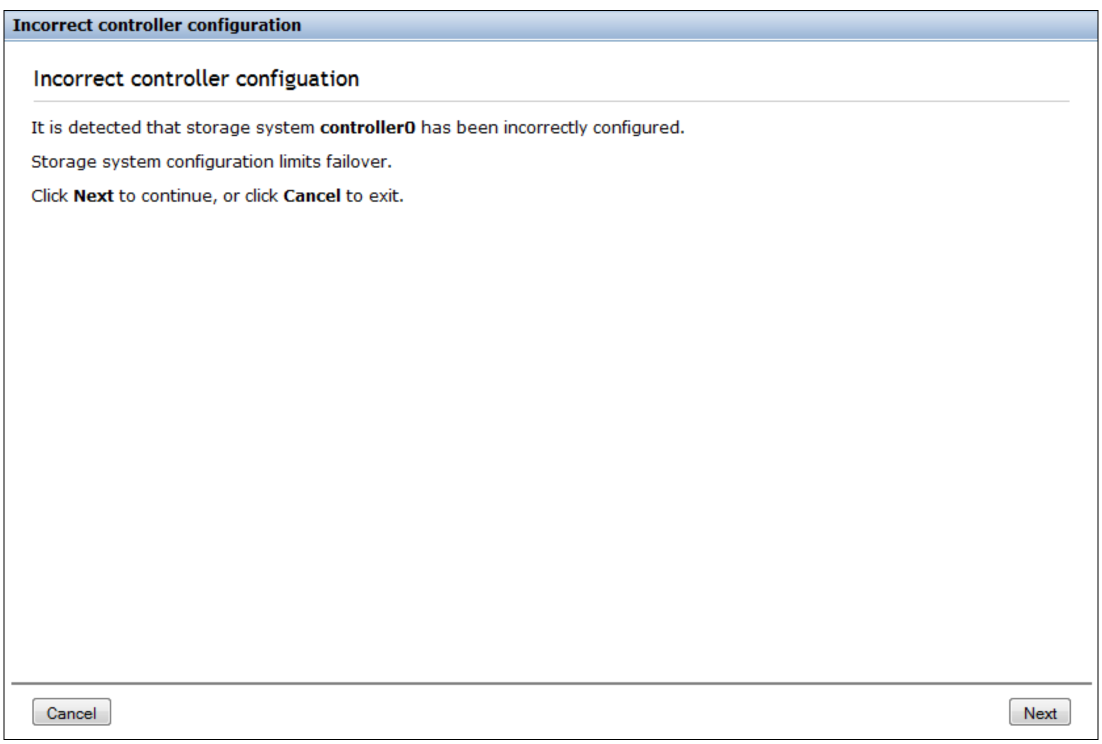


Figure 5-23 Error code 1625 Fix Procedure Step 1

Figure 5-24 shows the next step of the fix procedure. The fix procedure is informing about the error and what may have caused it.

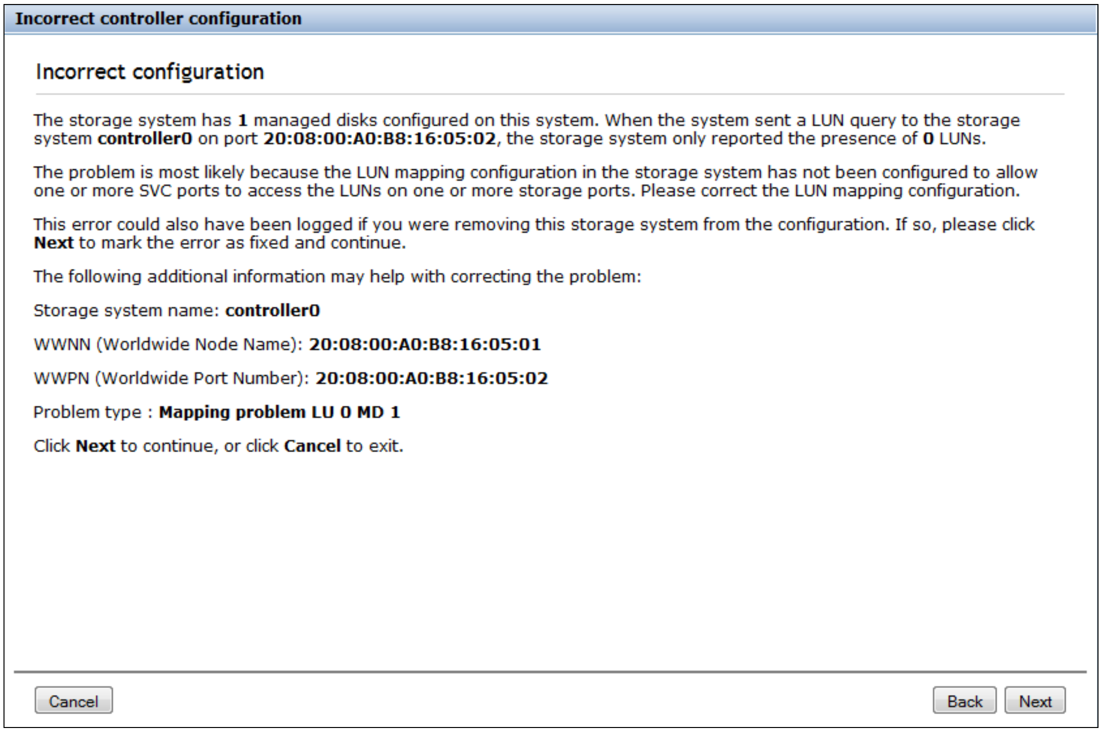


Figure 5-24 Error code 1625 Fix Procedure Step 2

Figure 5-25 on page 196 shows the third step of the fix procedure where we mark the error as fixed where after the IBM Flex System V7000 Storage Node starts a rescan of the system to see if the error persist.

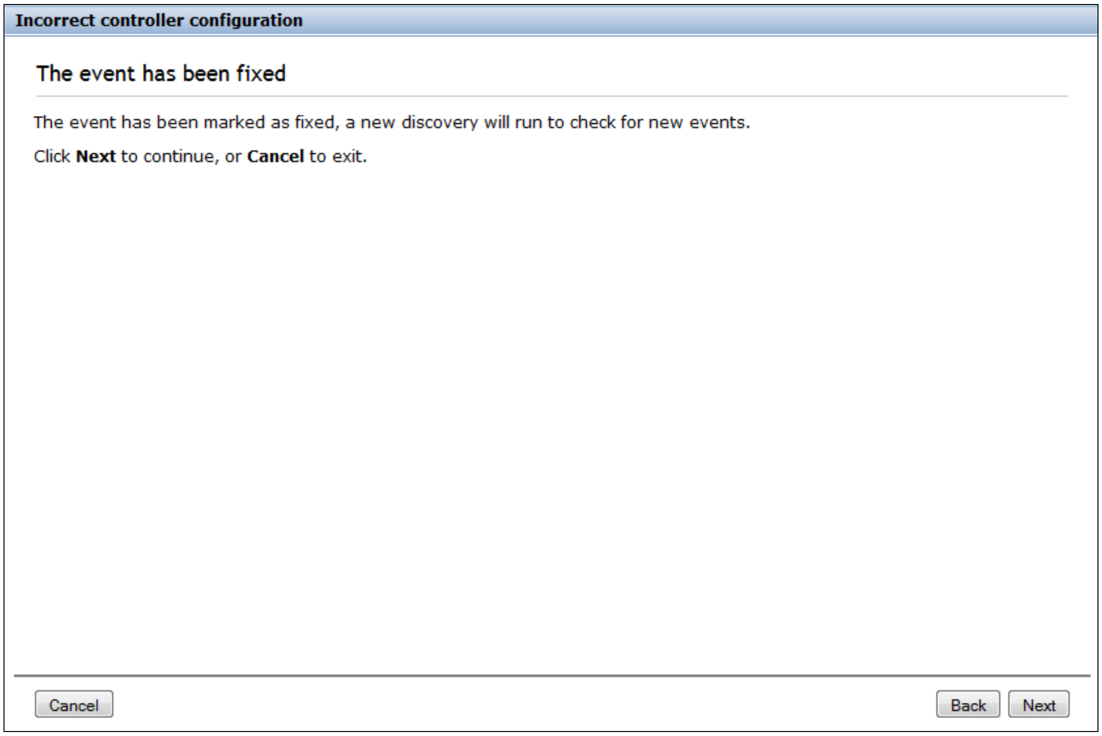


Figure 5-25 Error code 1625 Fix Procedure Step 3

Figure 5-26 shows that the rescan has begun.

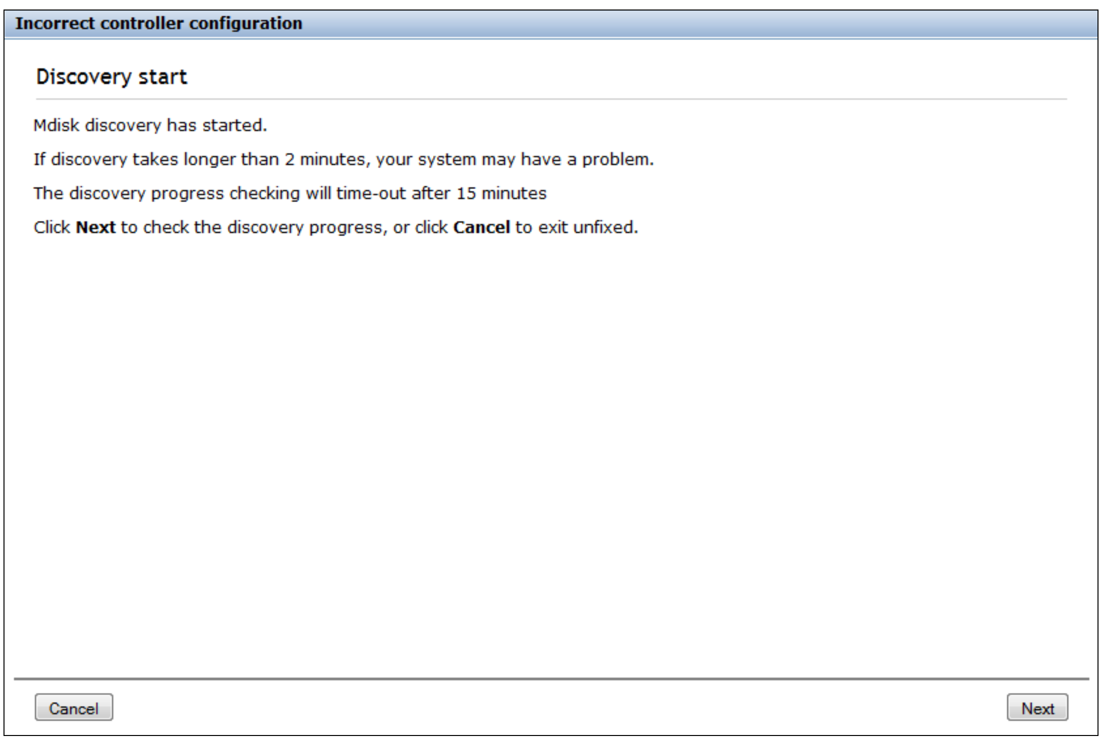


Figure 5-26 Error code 1625 Fix Procedure Step 4

Figure 5-27 on page 197 shows that the rescan has finished.

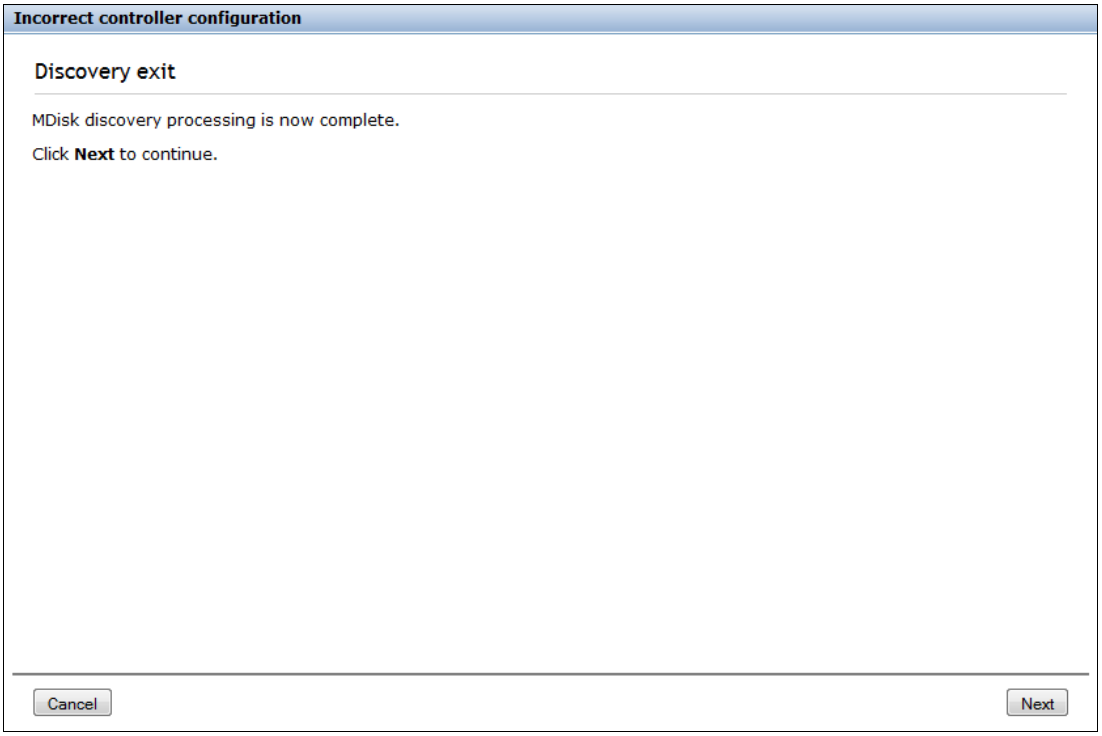


Figure 5-27 Error code 1625 Fix Procedure Step 5

Figure 5-28 shows that the rescan has finished and that the error condition has been resolved.

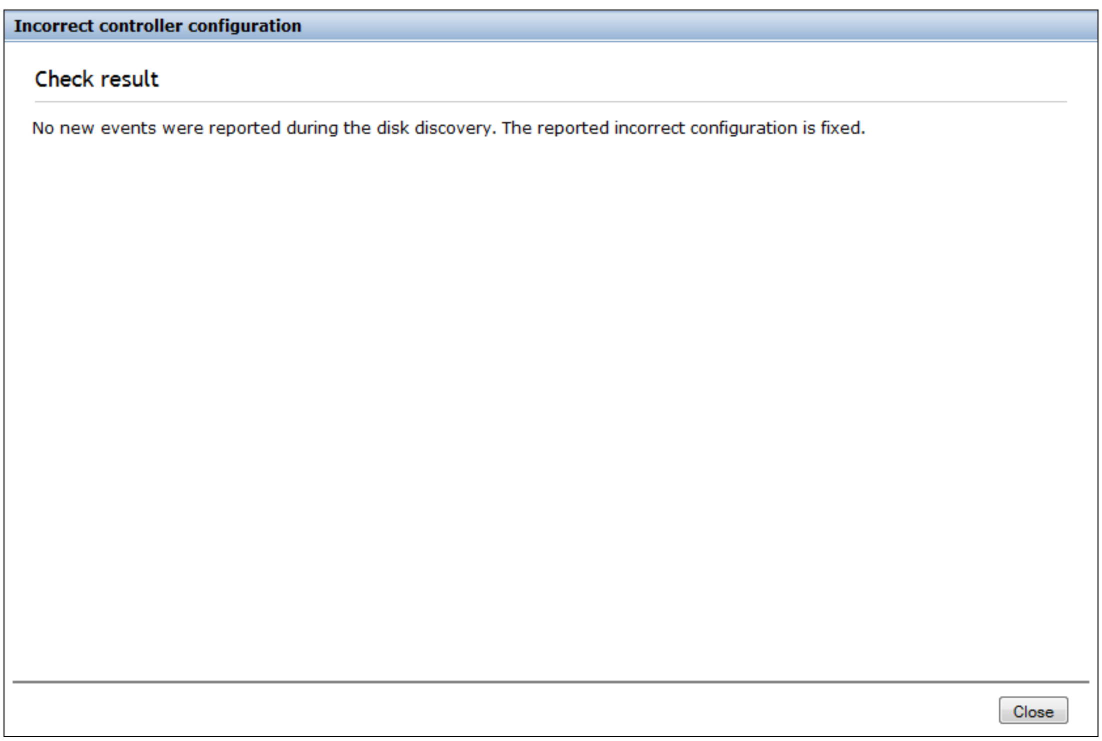


Figure 5-28 Error code 1625 Fix Procedure Step 6

The error 1625 in the above example was caused because of one of our Control Enclosure controllers had been temporarily removed from the system.

Figure 5-29 shows that few more errors remain as unresolved. These should be fixed as the one we just resolved in the example above.

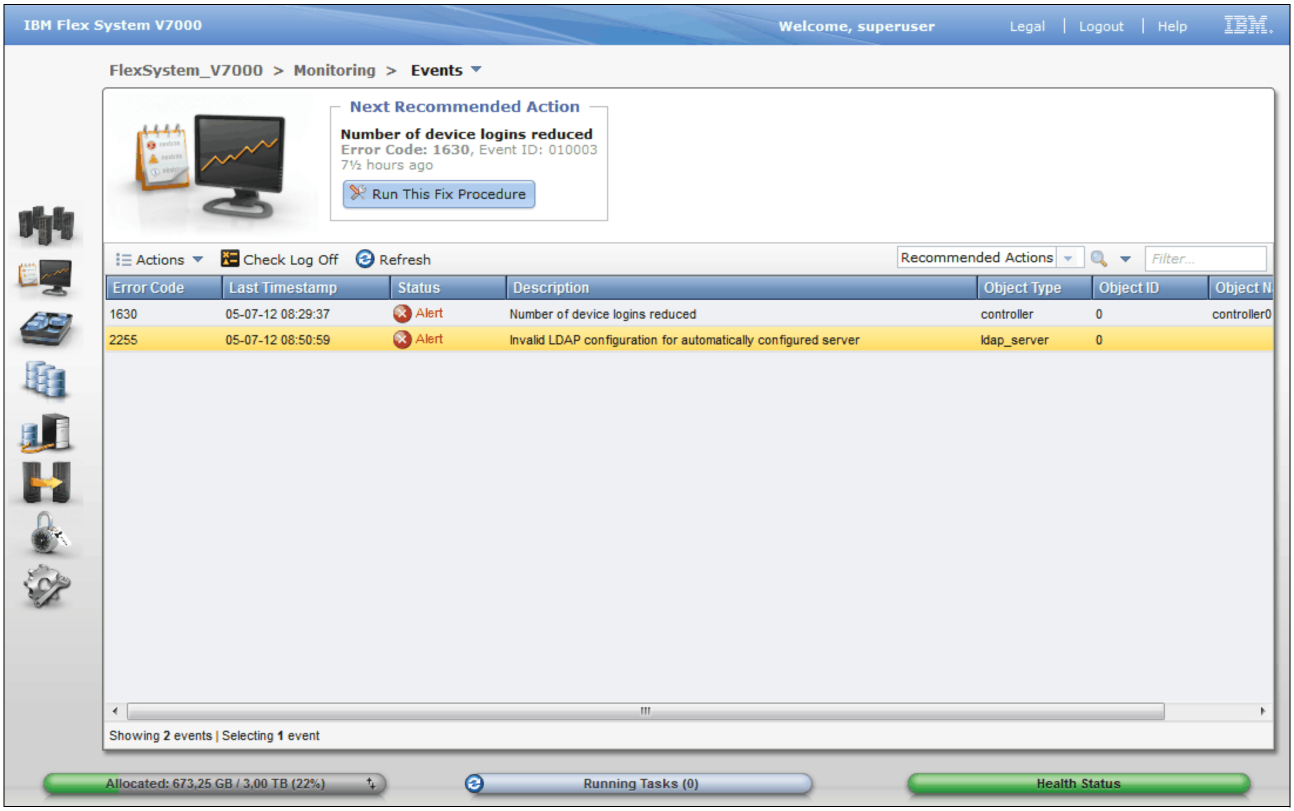


Figure 5-29 Error code 1625 Fix Procedure - Problem fixed

5.3.3 Monitoring Performance menu

The IBM Flex System V7000 Storage Node has a Performance menu which gives a first impression of how the system is performing. The performance chart is split into four sections:

- CPU Utilization - subdivided into:
  - CPU time used for the system
  - CPU time used to handle compression
- Interfaces - subdivided into:
  - FC - Fibre Channel traffic towards hosts
  - iSCSI - shows iSCSI traffic towards hosts
  - SAS - traffic towards the disk backend
- Volumes - traffic towards hosts that has mapped volumes - subdivided into:
  - Reads in MegaByte per second (MBps)
  - Writes in MegaByte per second (MBps)
  - Red latency in milli seconds (ms)
  - Write latency in milli seconds (ms)
- MDisks - traffic towards the disk backend - subdivided into:
  - Reads in MegaByte per second (MBps)

- Writes in MegaByte per second (MBps)
- Red latency in milli seconds (ms)
- Write latency in milli seconds (ms)

Figure 5-30 shows the Performance screen of IBM Flex System V7000 Storage Node.

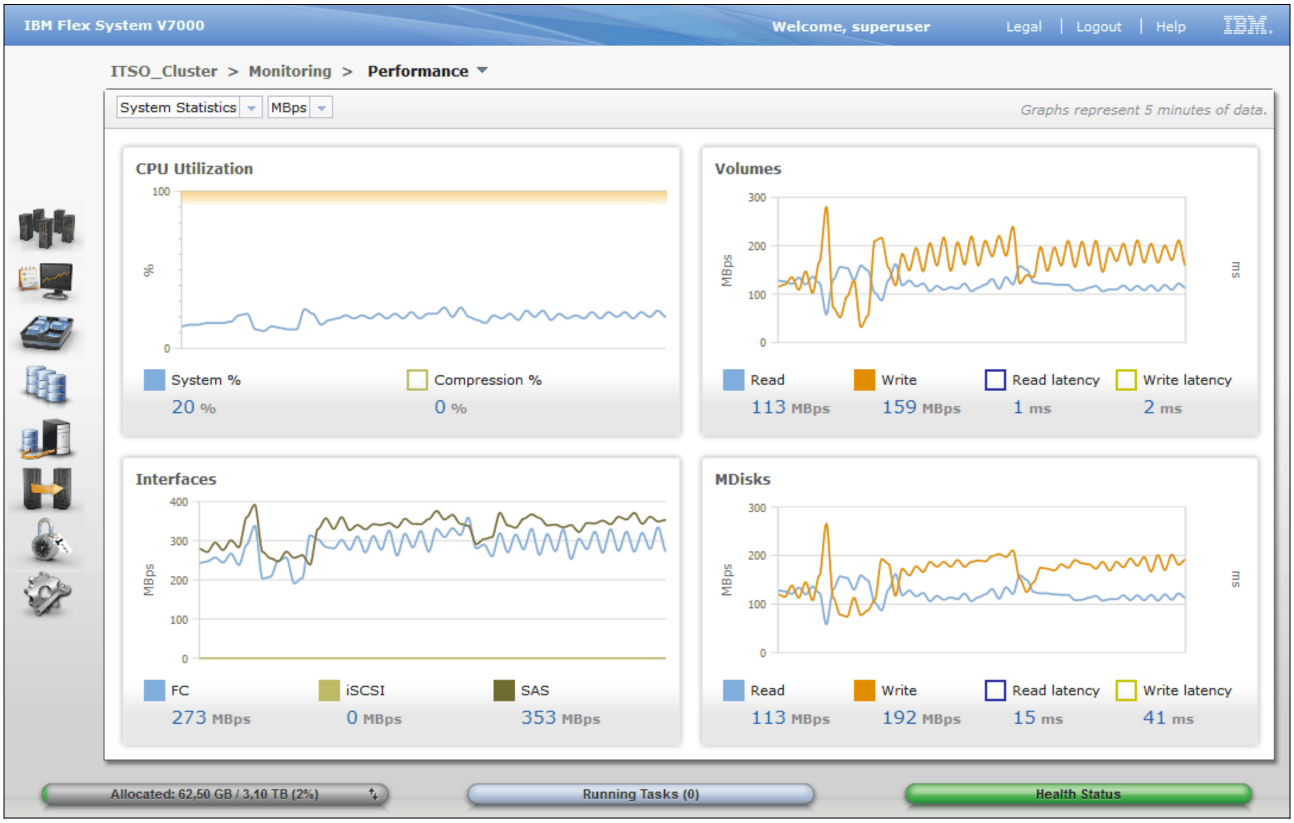


Figure 5-30 Monitoring Performance

If there is a need for additional performance monitors the most optimal tool is Tivoli Storage Productivity Center from where you can manage performance and connectivity from the host file system to the physical disk, including in-depth performance monitoring and analysis of storage area network (SAN) fabric.

For more information about Tivoli Storage Productivity Center refer to:

`http://www.ibm.com/systems/storage/software/center/index.html`

5.4 Pools menu

The storage pools are the pools of storage extents that are used to build the virtual volumes accessed by servers and used by the file server to form file systems.

The Pools menu contain different views of the IBM Flex System V7000 Storage Node Volumes, MDisks, Internal storage, External storage as well as the System Migration functionality.

The When you hover the cursor over the Pools function icon, a set of menu options appears (Figure 5-31 on page 200).

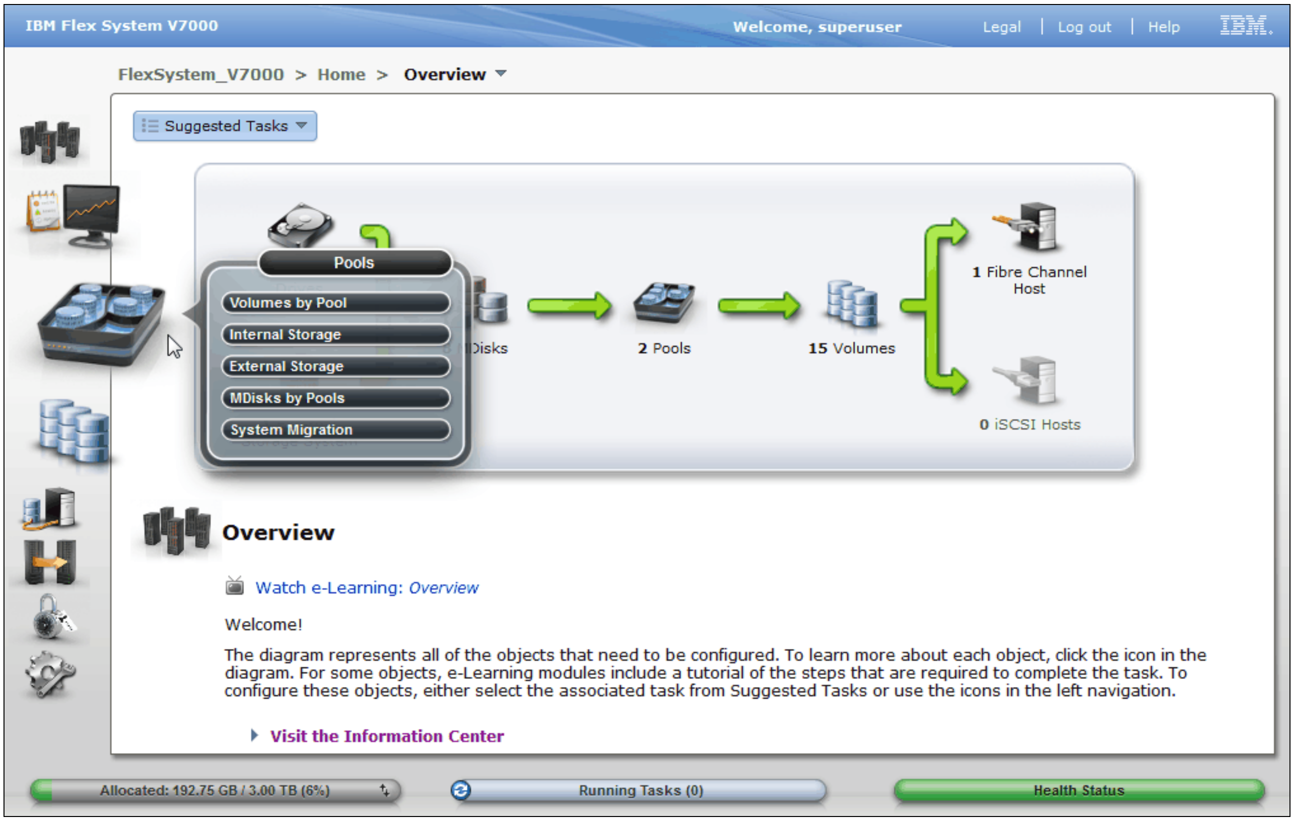


Figure 5-31 Navigate to the Pools menu

Volumes by Pool Menu is discussed in 5.4.1, “Volumes by Pool menu” on page 201.

5.4.1 Volumes by Pool menu

Figure 5-32 shows the Volumes by Pool window, from where you can create or delete storage pools as well as perform actions on volumes like mapping to host or unmapping from host. Also migration actions can be performed from the Volumes by Pool menu

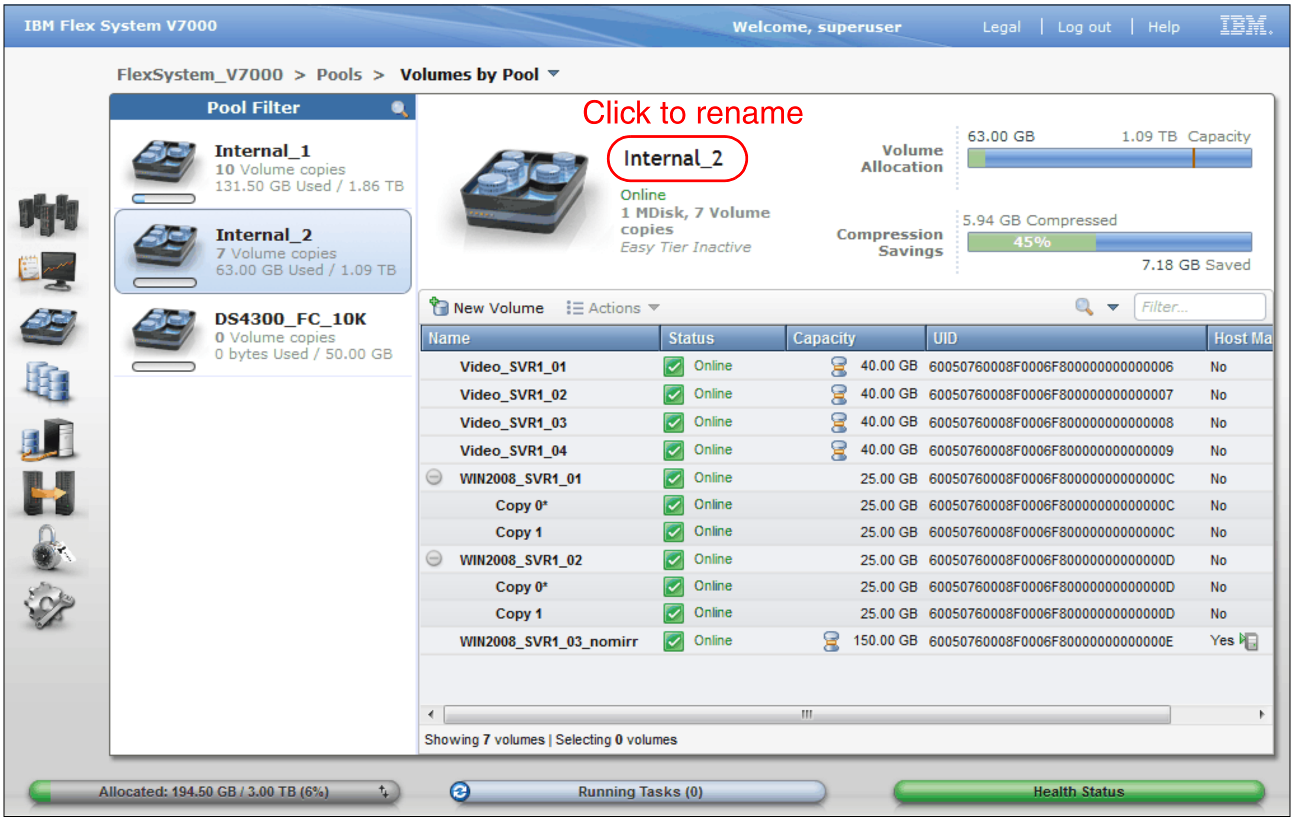


Figure 5-32 Volumes by Pools menu

As with the storage controllers, you may rename the MDisk pools to reflect where the disks are located.

5.4.2 Internal Storage menu

The IBM Flex System V7000 Storage Node has internal disk drives in the enclosures. The Internal Storage menu option displays and manages these drives.

Clicking the **Internal Storage** option opens the window shown in Figure 5-33. From this window, you can place internal disk drives into storage pools. This window also provides the option to display the internal drives based on their capacity and speed.

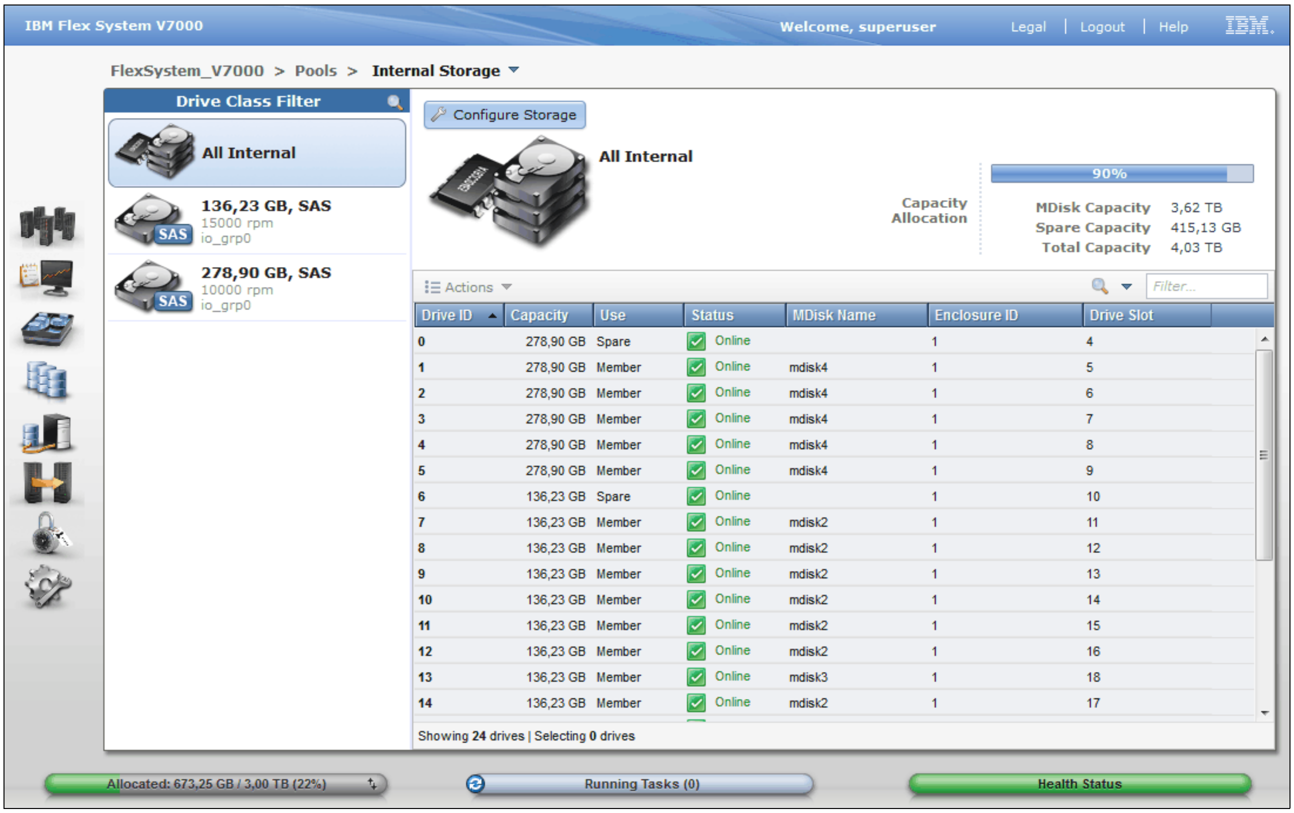


Figure 5-33 Internal storage window

For more information about how to configure and manage internal storage refer Chapter 8, “Storage pools” on page 295.

5.4.3 External Storage menu

The IBM Flex System V7000 Storage Node can also manage external storage subsystems using the SAN or FCoE connections. If any are attached, they are managed in this option.

Clicking the **External Storage** option opens the window shown in Figure 5-34. This window shows any virtual external disk systems in that IBM Flex System V7000 Storage Node. From this window, you can add MDisks to existing pools, import them as image mode volumes, or rename them. If you highlight a MDisk, you can also display any dependent volumes.

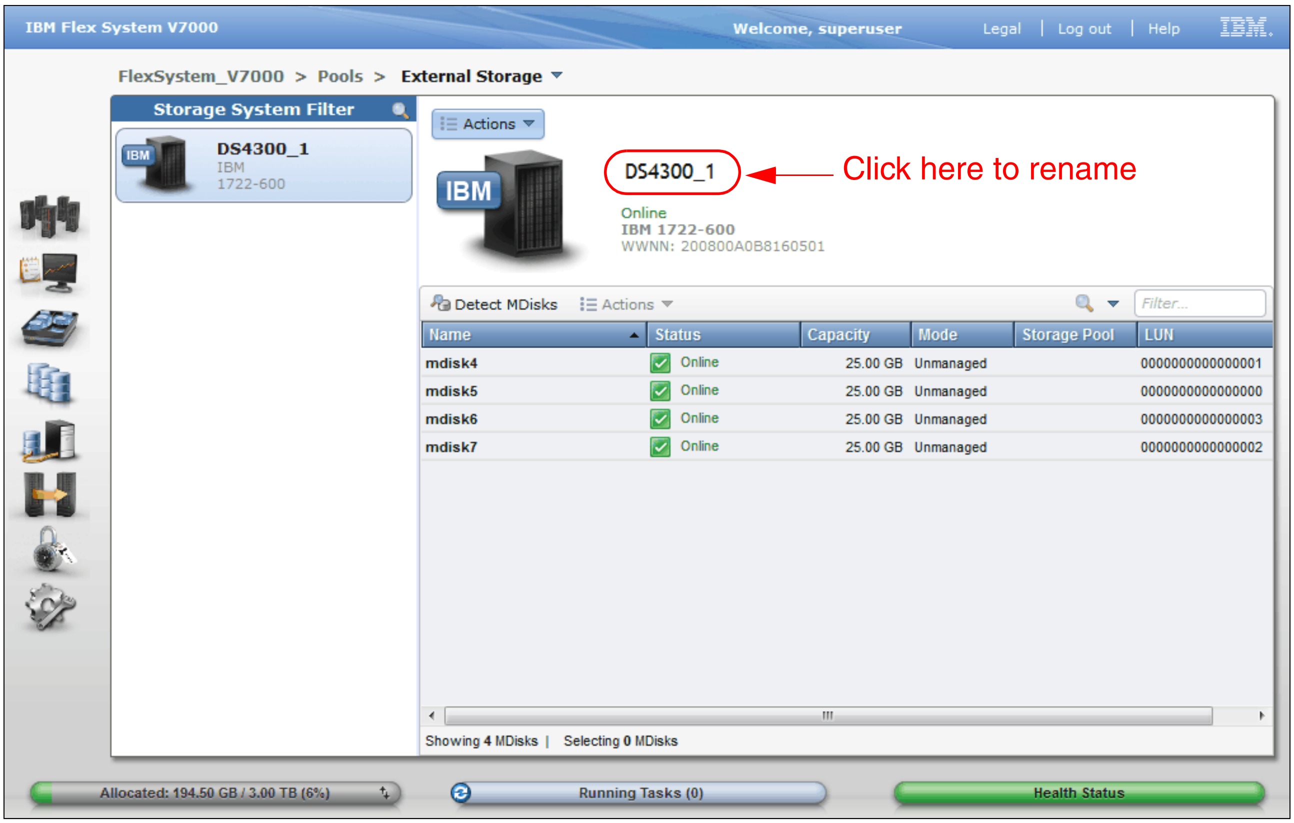


Figure 5-34 External Storage systems menu

By default the external storage controllers are named controller0, controller1 etc. It is good practice to rename these controllers so that the name reflect the hostname of the storage system. This is done by clicking the controller name in the right side of the window. You may also rename the MDisks to reflect LUN names from the storage system presenting the disks to the IBM Flex System V7000 Storage Node.

For more information about how to virtualize external storage refer to Chapter 7, “Migration wizard” on page 265.

5.4.4 System Migration

System Migration is used to virtualize and migrate existing external storage disks to the IBM Flex System V7000 Storage Node. The System Migration tools are described in detail in Chapter 7, “Migration wizard” on page 265.

5.4.5 MDisks by Pools menu

Figure 5-35 shows the MDisks that are available to the IBM Flex System V7000 Storage Node system. The MDisks show whether they are *managed*, in which case the storage pool is displayed, or whether they are *unmanaged*, in which case the storage pools can be added to a new pool.

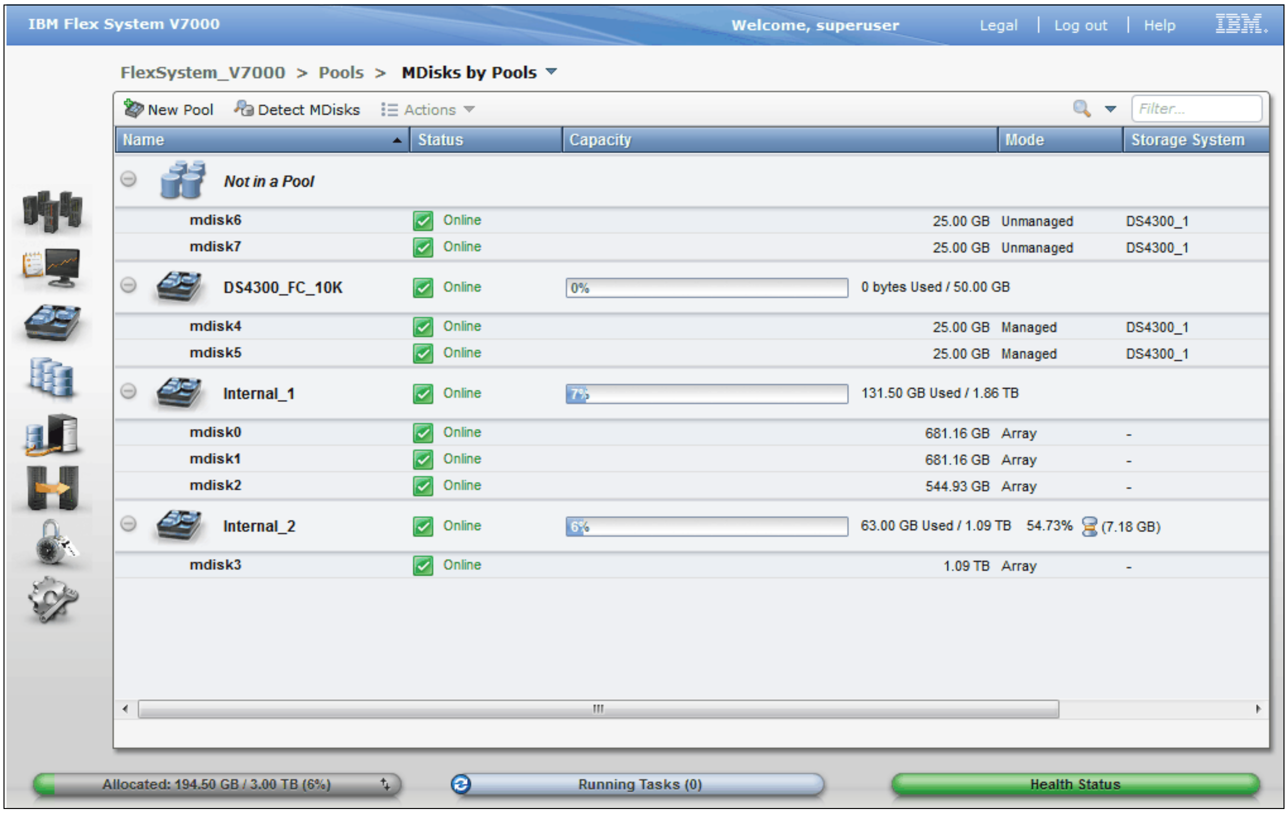


Figure 5-35 MDisks by Pools menu

By right-clicking the display bar (Figure 5-36), you can choose to change the fields that are displayed. Select the items you want to be displayed.

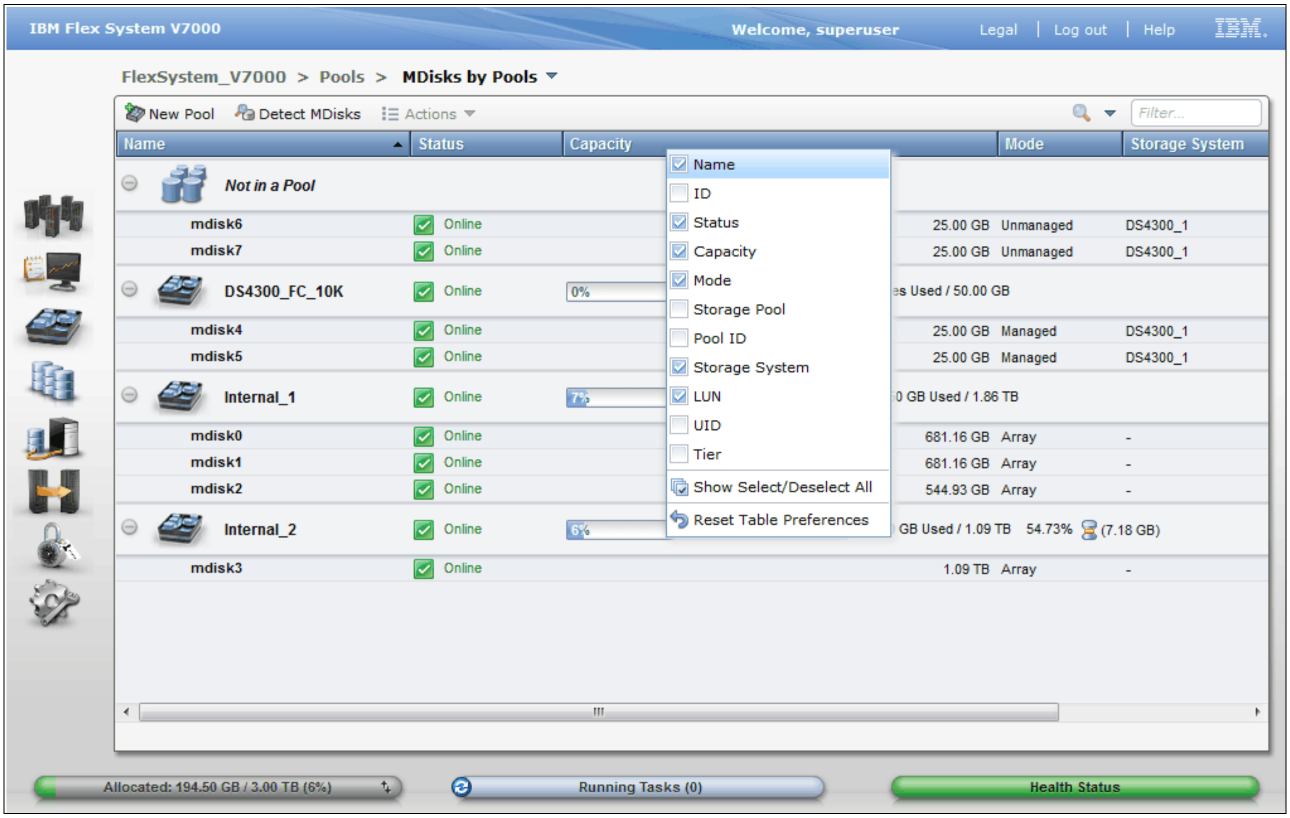


Figure 5-36 Display additional fields

From this window, you can choose to either use the options of the **Actions** button or you can choose to highlight the particular MDisks that you require, right-click, and access the same options (Figure 5-37).

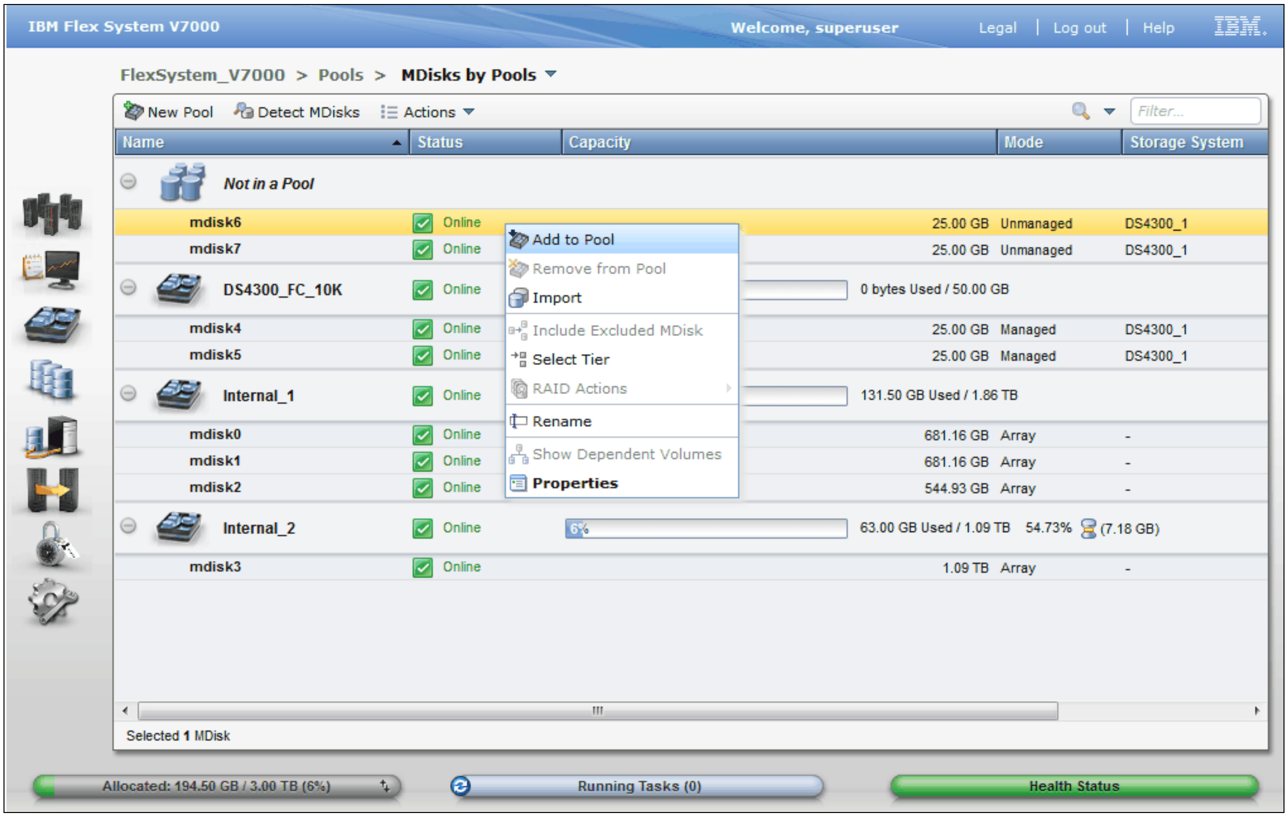


Figure 5-37 Actions for a single MDisk from the MDisks by Pools menu

5.5 Volumes menu

The volumes are built from extents in the pools, and are presented to hosts as external disks. In this section, we describe the Volumes menu and its options. When you hover the cursor over the Volumes function icon, the menu shown in Figure 5-38 opens.

Figure 5-38 Navigate to the Volumes menu

5.5.1 The Volumes menu

Clicking the **Volumes** menu opens the window shown in Figure 5-39. From here you can perform tasks on the volumes, such as shrink or enlarge them, map them to a host, or migrate a volume.

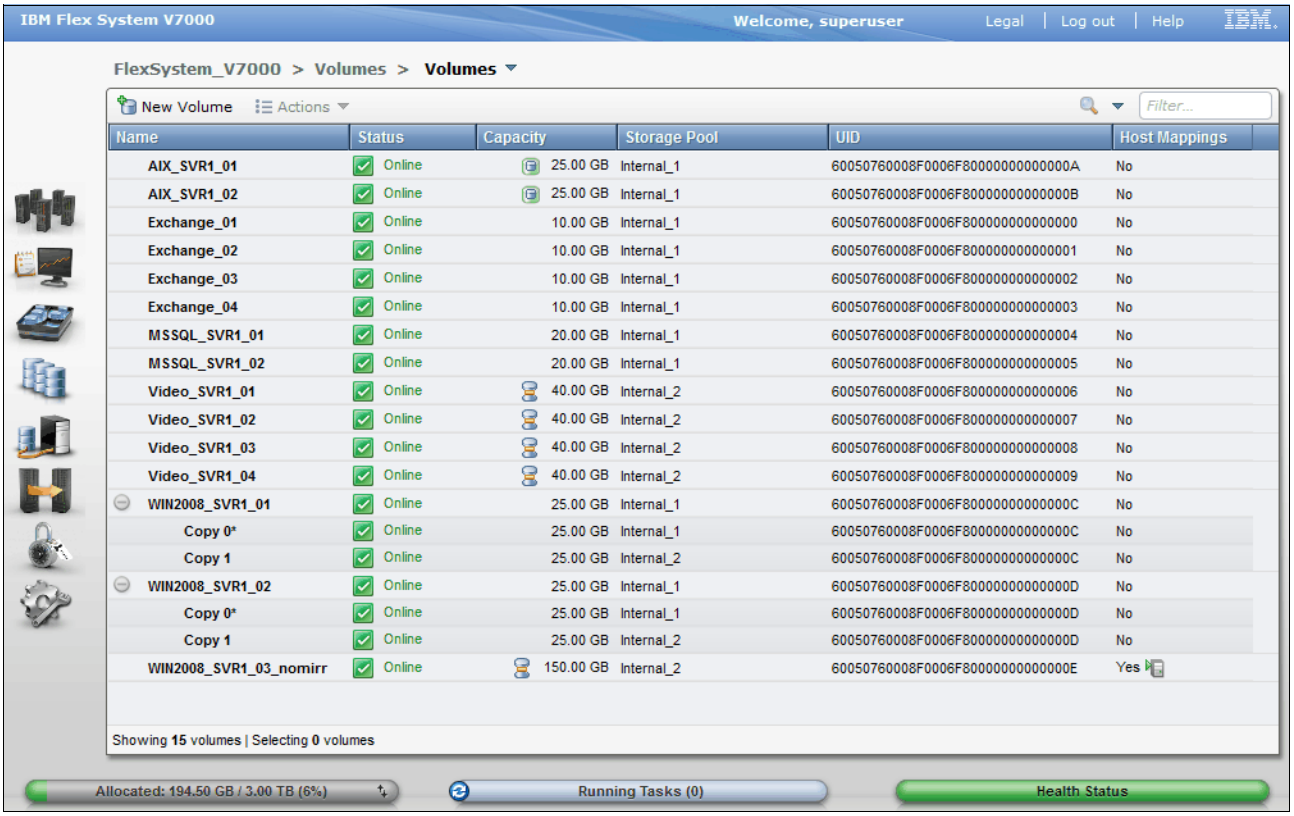


Figure 5-39 Volumes window that shows all volumes

From this menu, you can perform various operations on the volumes. You can use the **Actions** button to access these operations, or you can right-click the Volume name, which opens list of operations that can be performed against the volume (Figure 5-40).

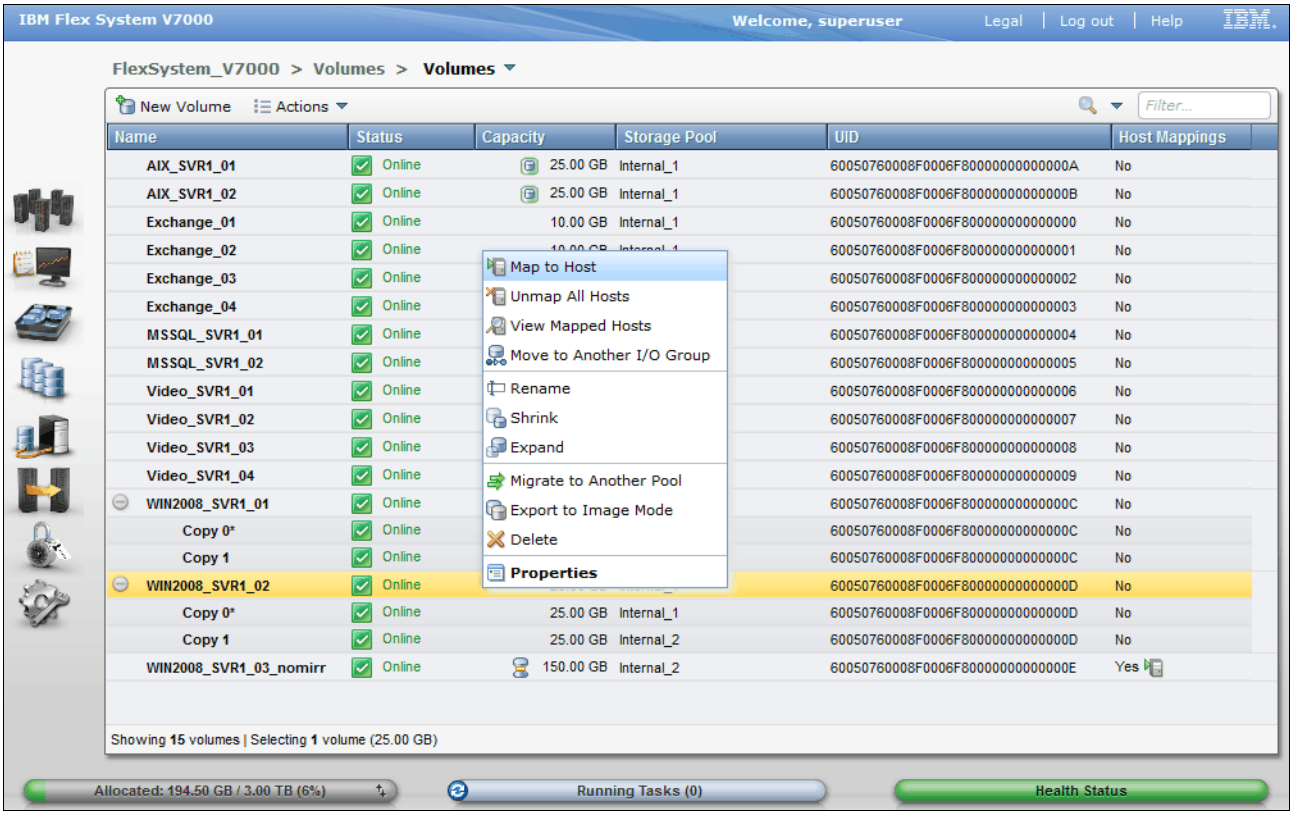


Figure 5-40 Volume operations

5.5.2 Volumes by Pool menu

Clicking the **Volumes by Pool** menu opens the window shown in Figure 5-41.

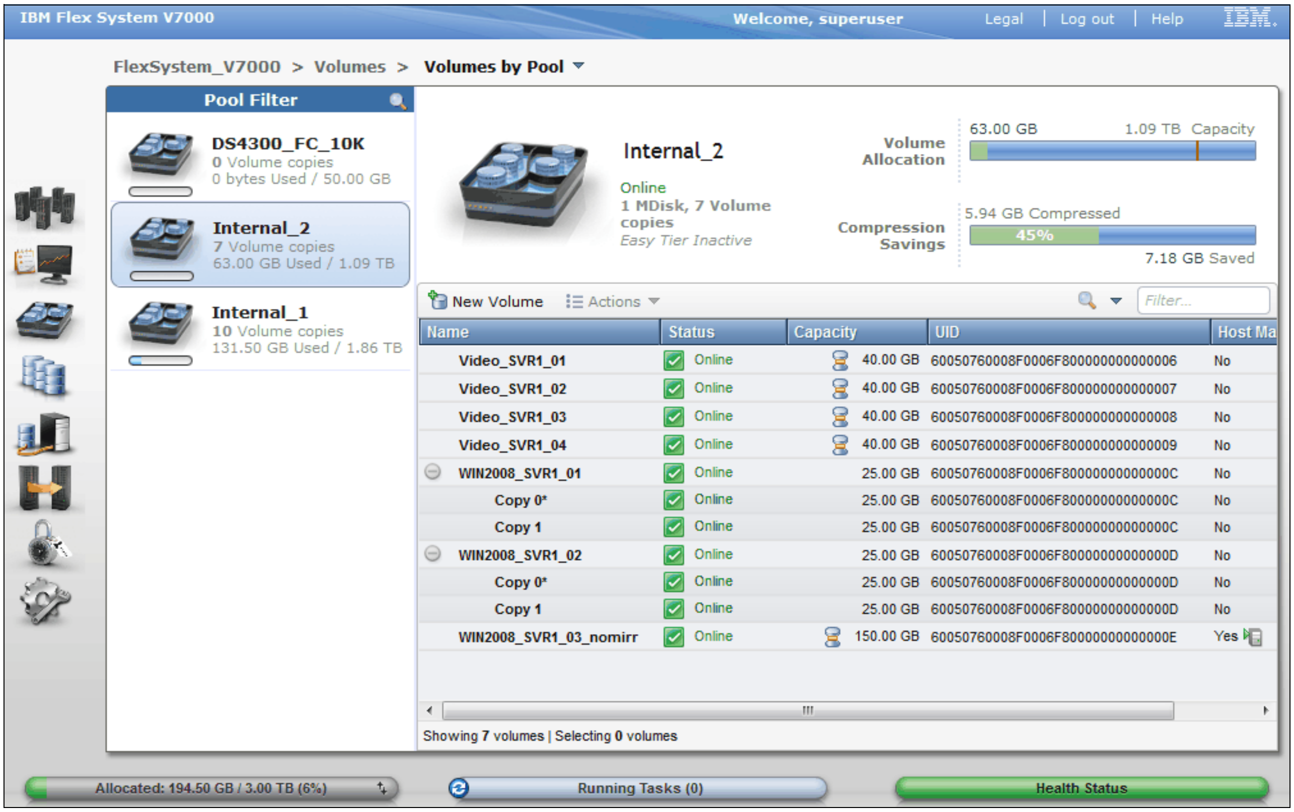


Figure 5-41 Volumes by Pool window

Similar to the previous window, you can either use the **Actions** button to access the menu operations or you can right-click the pool to display a list of valid commands (Figure 5-42).

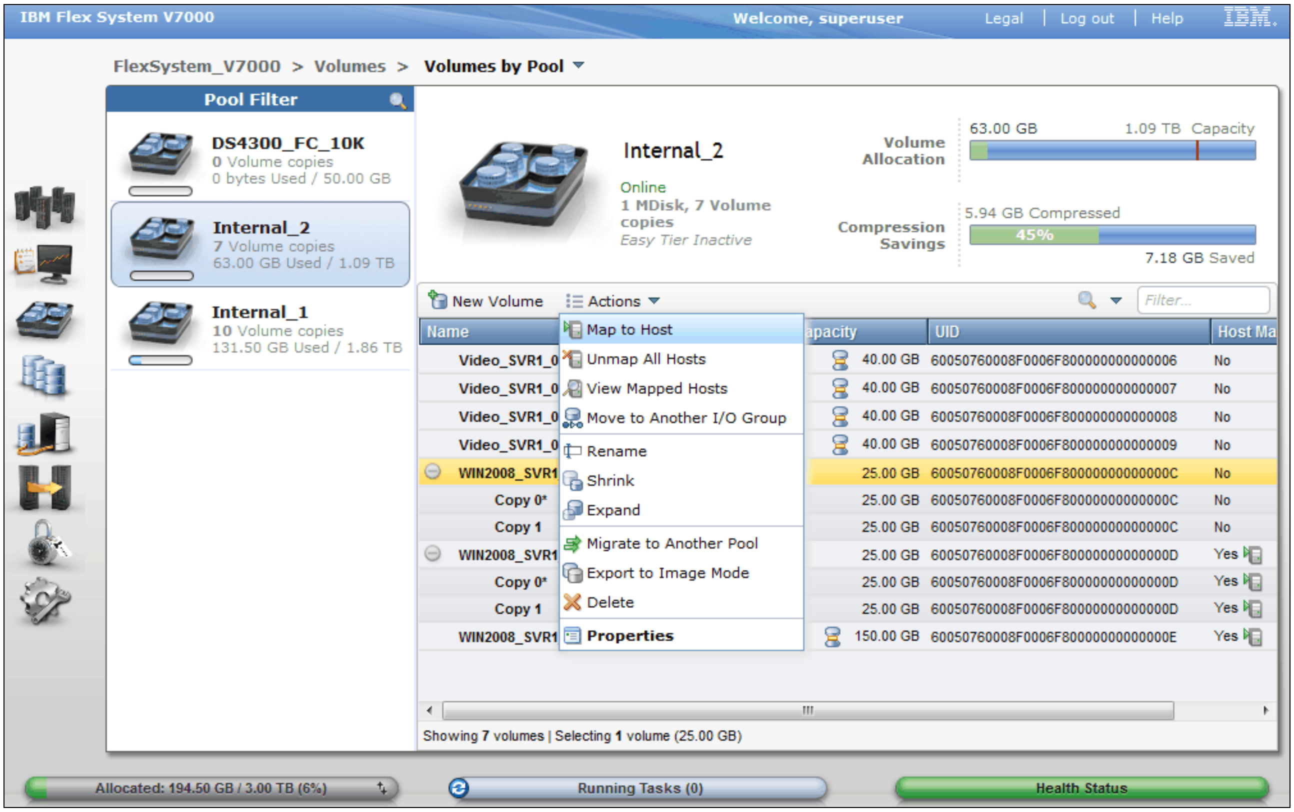


Figure 5-42 Commands for a single volume from the Volume by Pools menu

5.5.3 Volumes by Host menu

Clicking the **Volumes by Host** option opens the window shown in Figure 5-43 on page 212. This window shows the volumes that have been mapped to a given host.

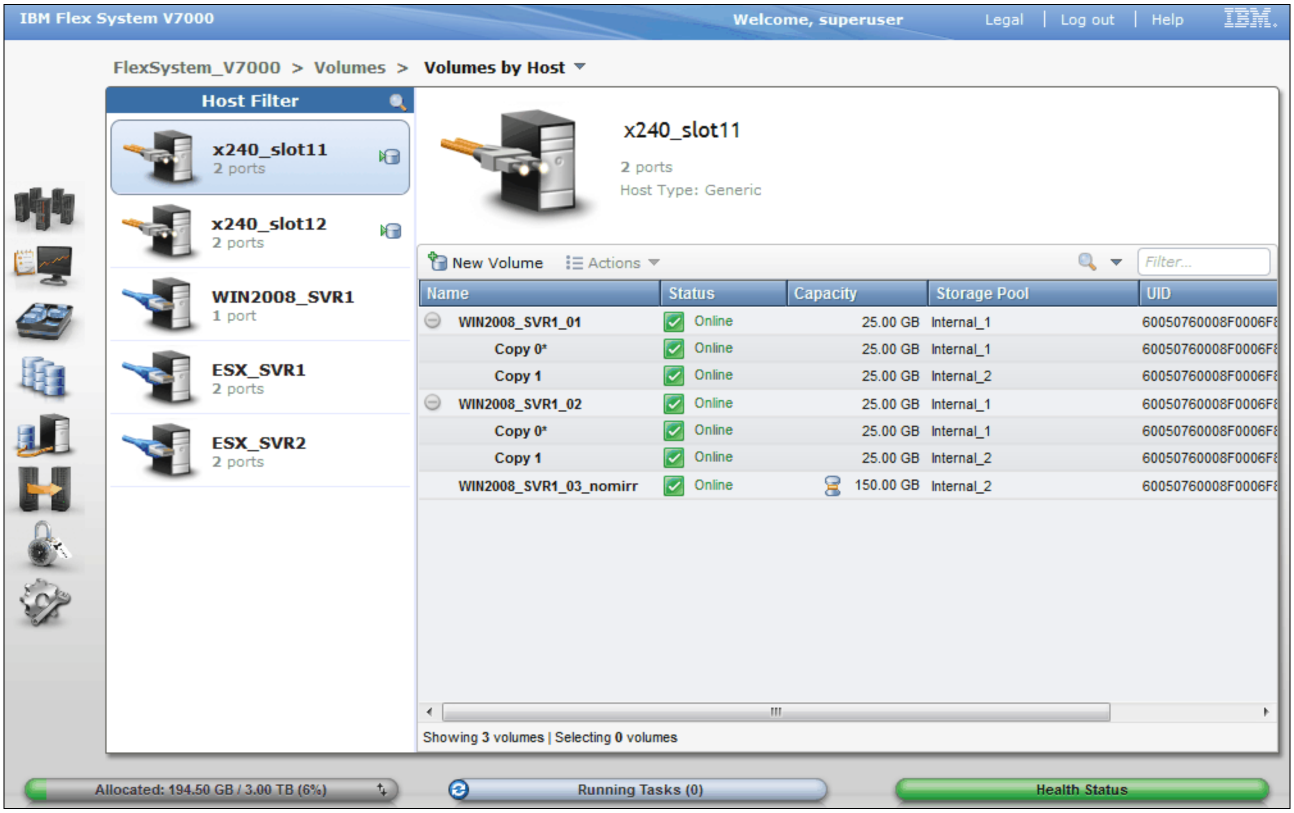


Figure 5-43 Volumes by Host window

You can use the **Actions** button or you can right-click the pool to show a list of valid commands (Figure 5-44 on page 213).

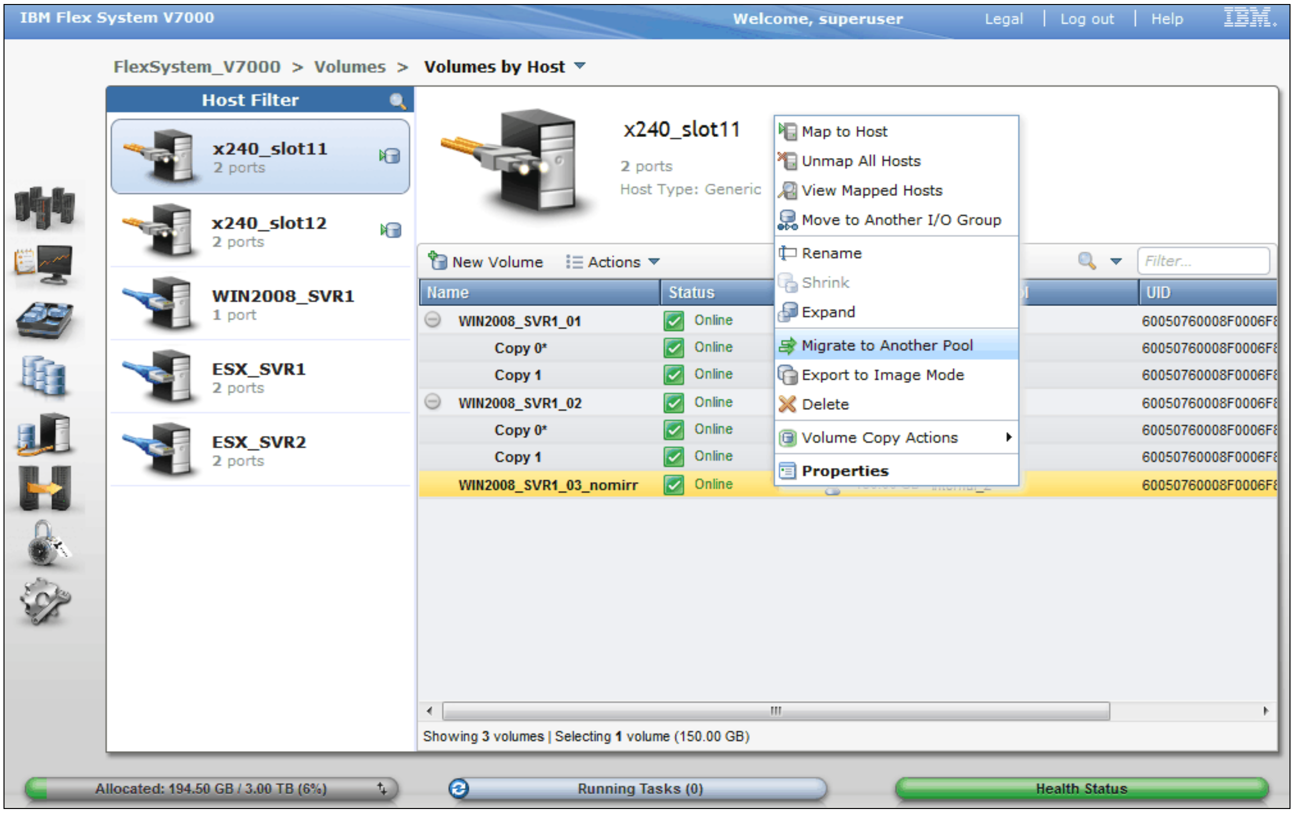


Figure 5-44 Commands for a single volume from the Volume by Host menu

By using the **Migrate to Another Pool** option the Volume can be migrated from one pool to another while still maintaining full service to the host that connects to the Volume.

5.6 Hosts menu

In this section, we describe the Hosts menu and its options. When you hover the cursor over the Host function icon, a menu opens (Figure 5-45).

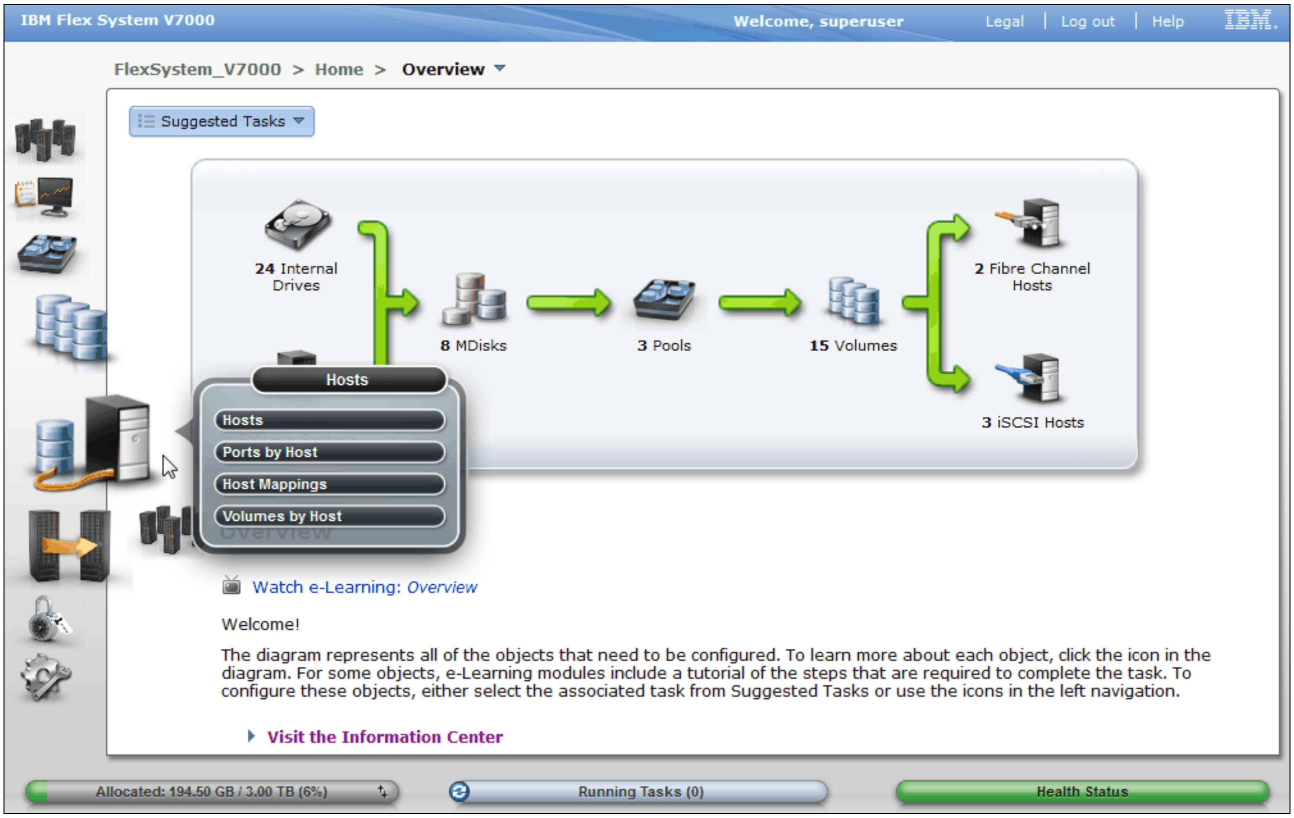


Figure 5-45 Navigate to the Hosts menu

5.6.1 The Hosts menu

Clicking **Hosts** opens the window shown in Figure 5-46. From here you can modify host mappings, unmap hosts, rename hosts, and create new hosts.

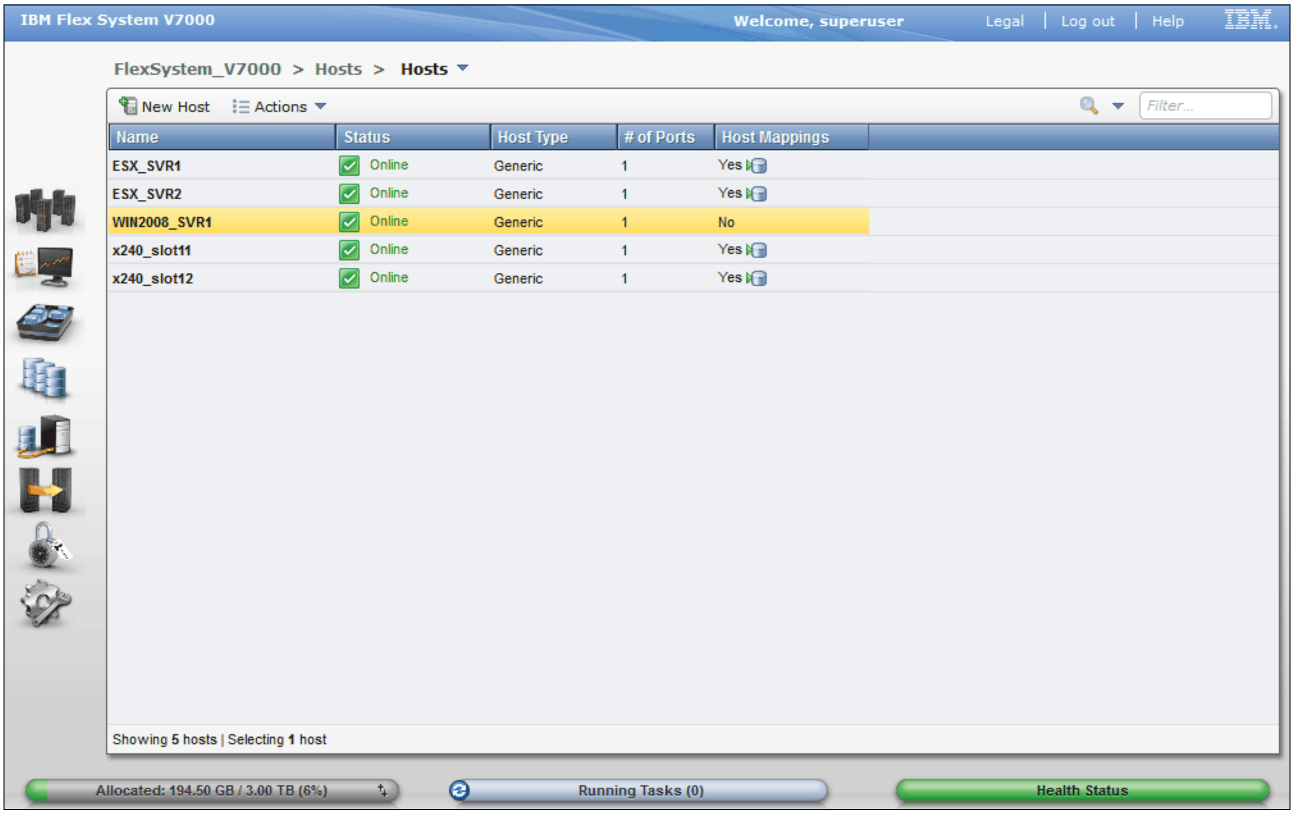


Figure 5-46 The Hosts window that shows all hosts

As with a number of other windows, you can use the command buttons or you can select a host and right-click it to access the commands (Figure 5-47).

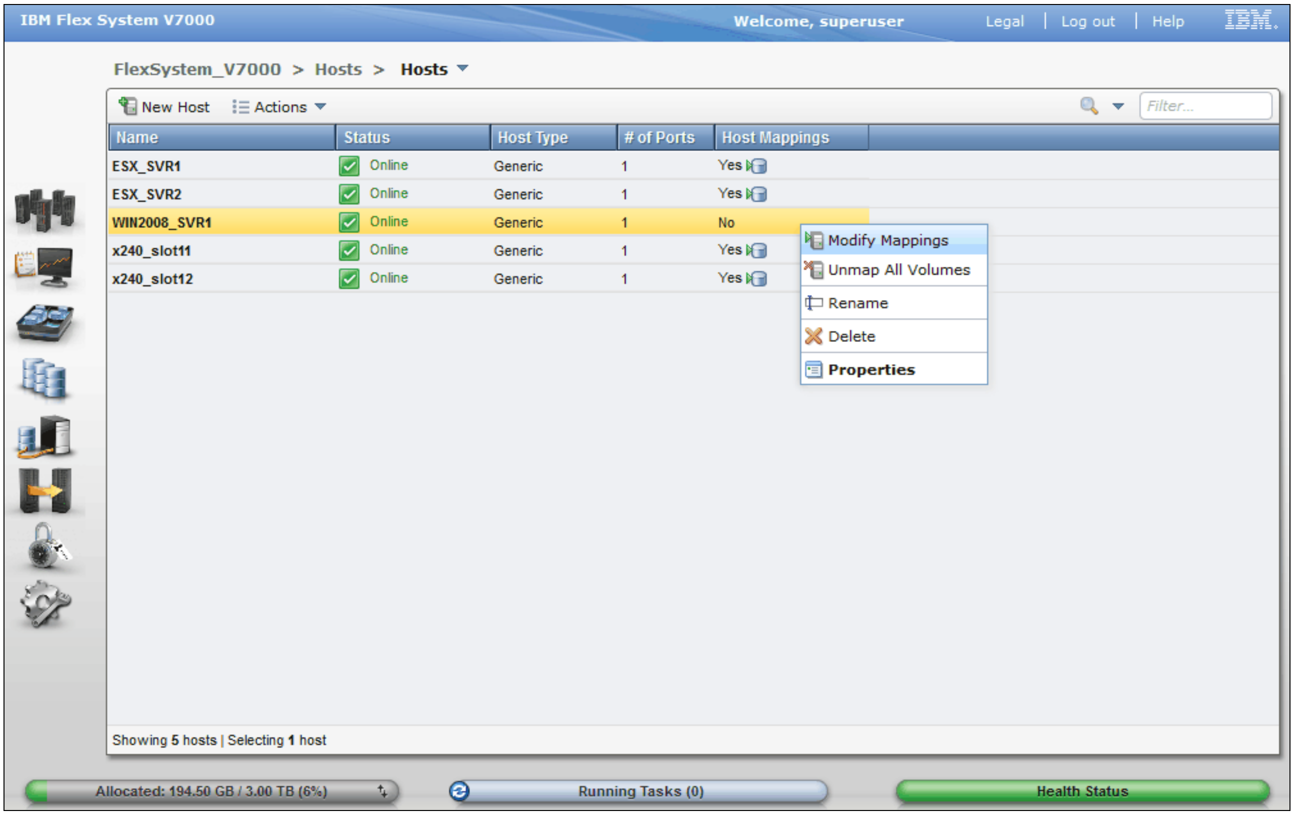


Figure 5-47 Actions for a single host from the Hosts menu

Figure 5-48 shows an example of Modify Mappings. The selected Windows server has no mappings and we select a Volume for to be mapped to the server. From this menu we can also map more volumes or we can unmap volumes already mapped.

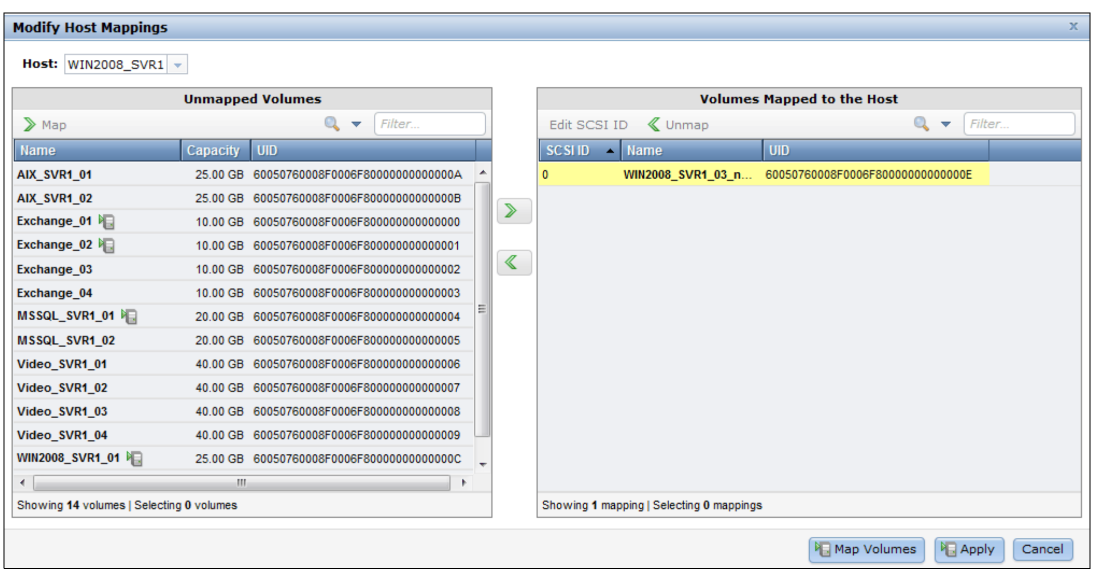


Figure 5-48 Modify Mappings

5.6.2 Ports by Host menu

Clicking **Ports by Hosts** opens the window shown in Figure 5-49. This window shows the Fibre Channel and iSCSI ports that are assigned to a particular host.

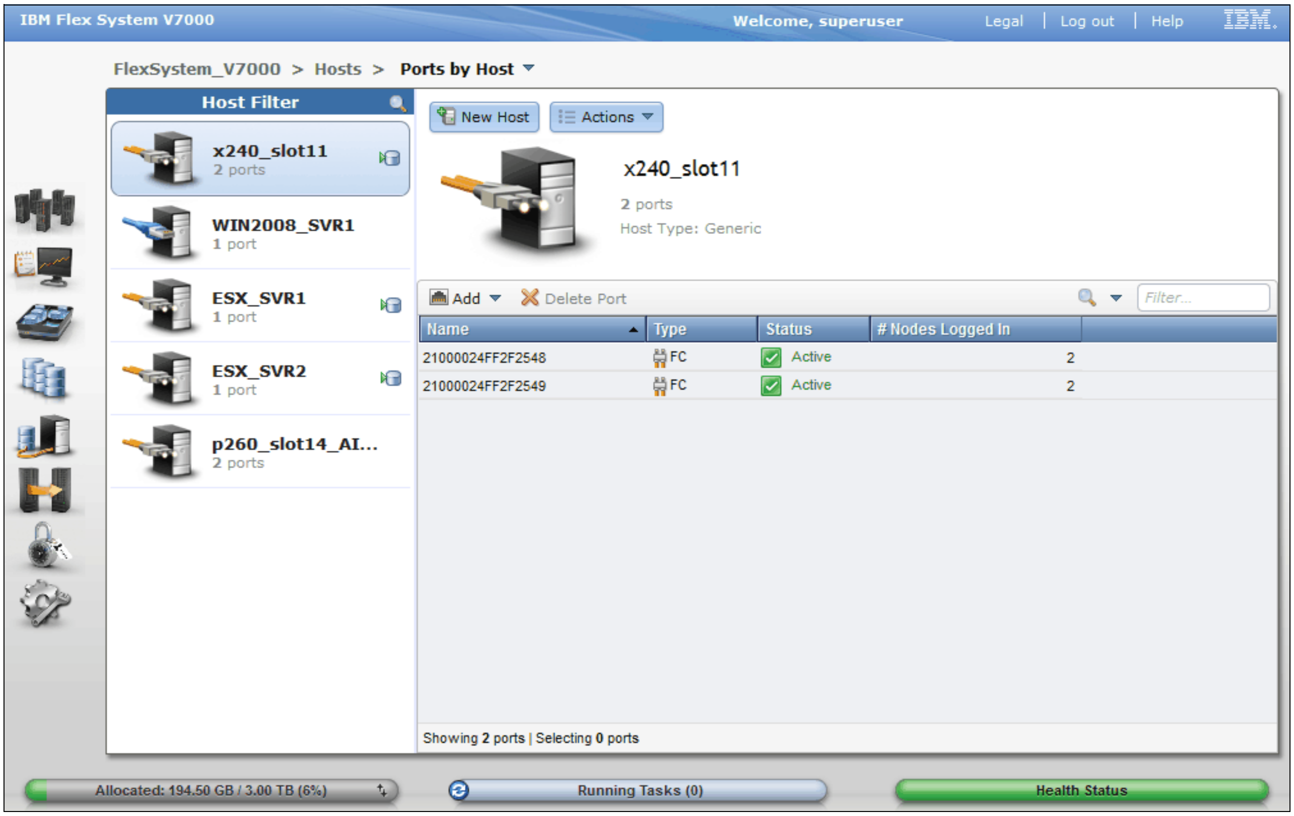


Figure 5-49 Ports by Host window

Clicking the **Actions** button allows you to modify the mappings, unmap volumes, rename hosts, and delete ports (Figure 5-50).

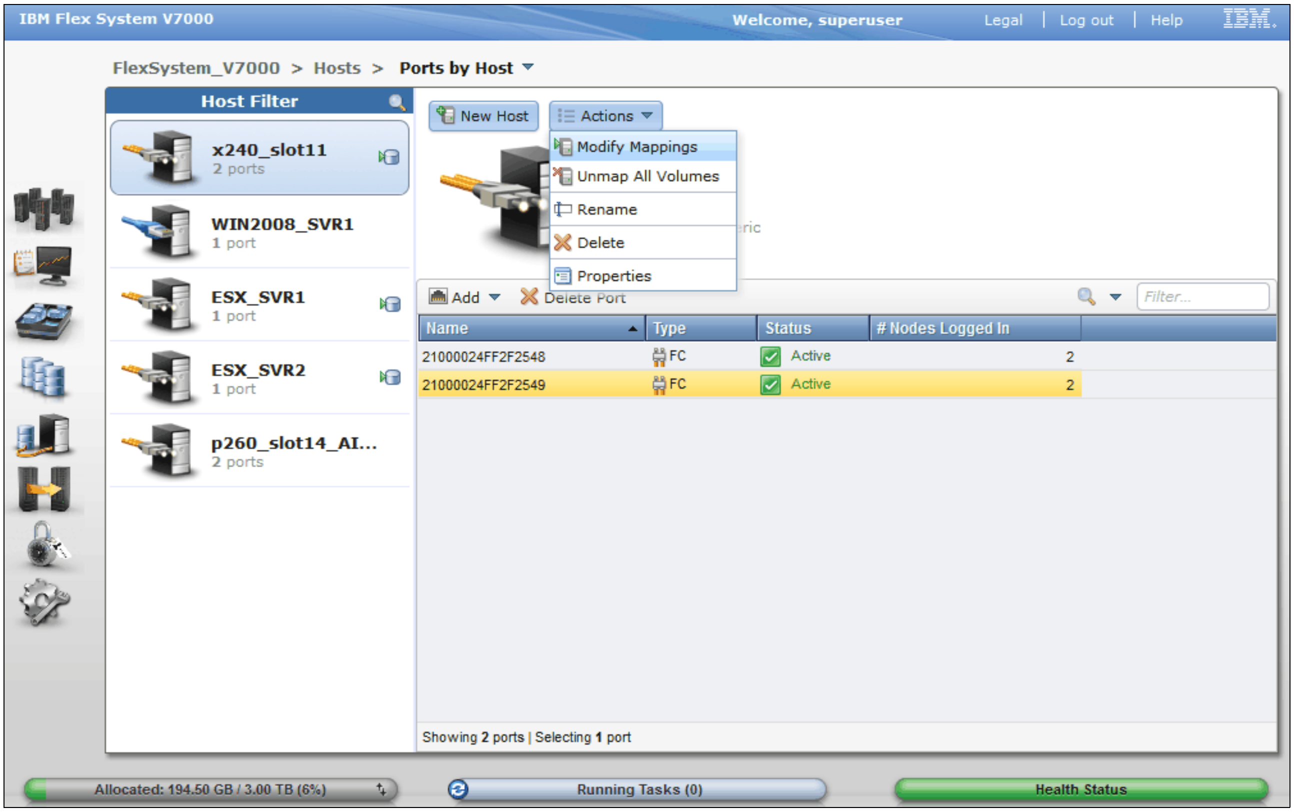


Figure 5-50 Actions for a single host from the Ports by Host window

5.6.3 Host Mappings menu

Clicking **Host Mappings** opens the window shown in Figure 5-51. This window shows the Host name, SCSI identifier, Volume name and the Volume identifier for all the mapped volumes.

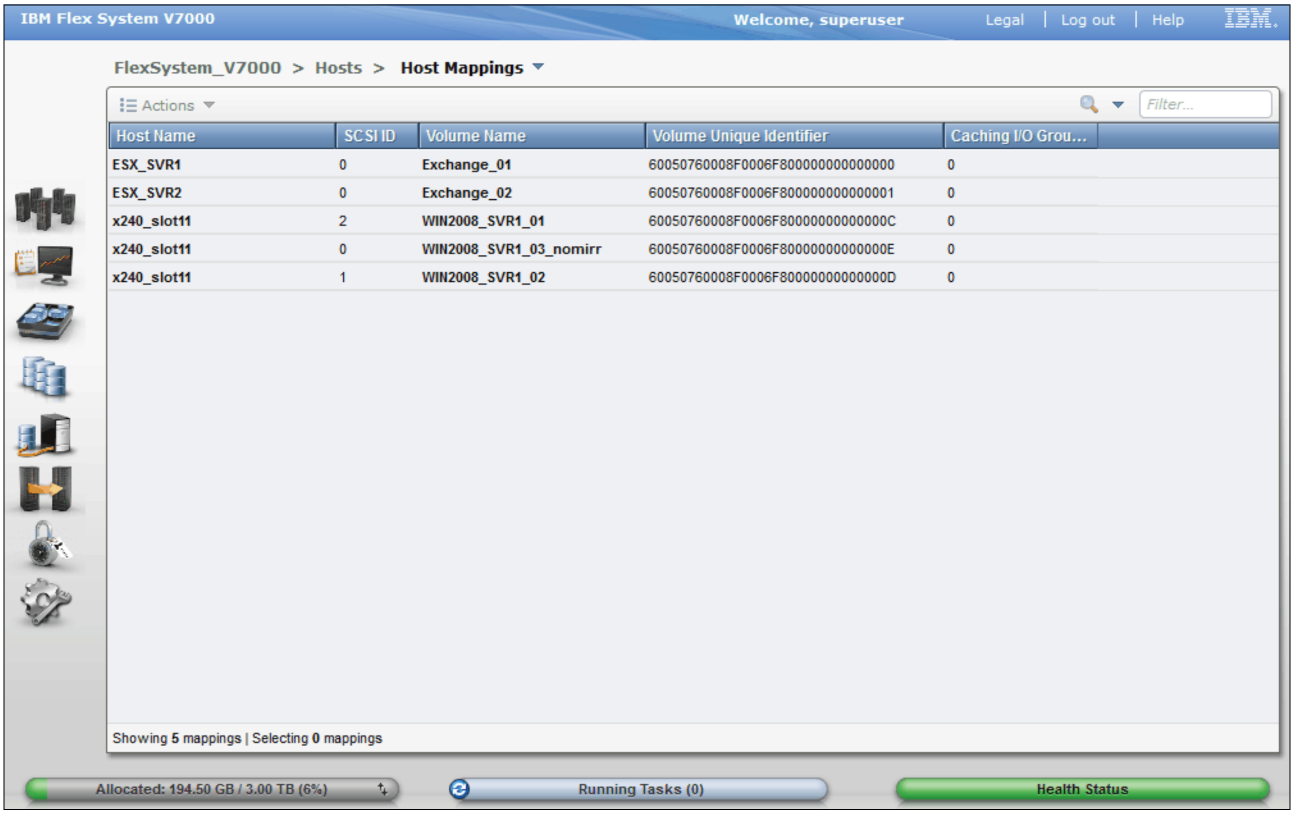


Figure 5-51 Host Mapping window

You can use the Actions buttons shown in Figure 5-52 or you can select a host and right-click it to access the commands.

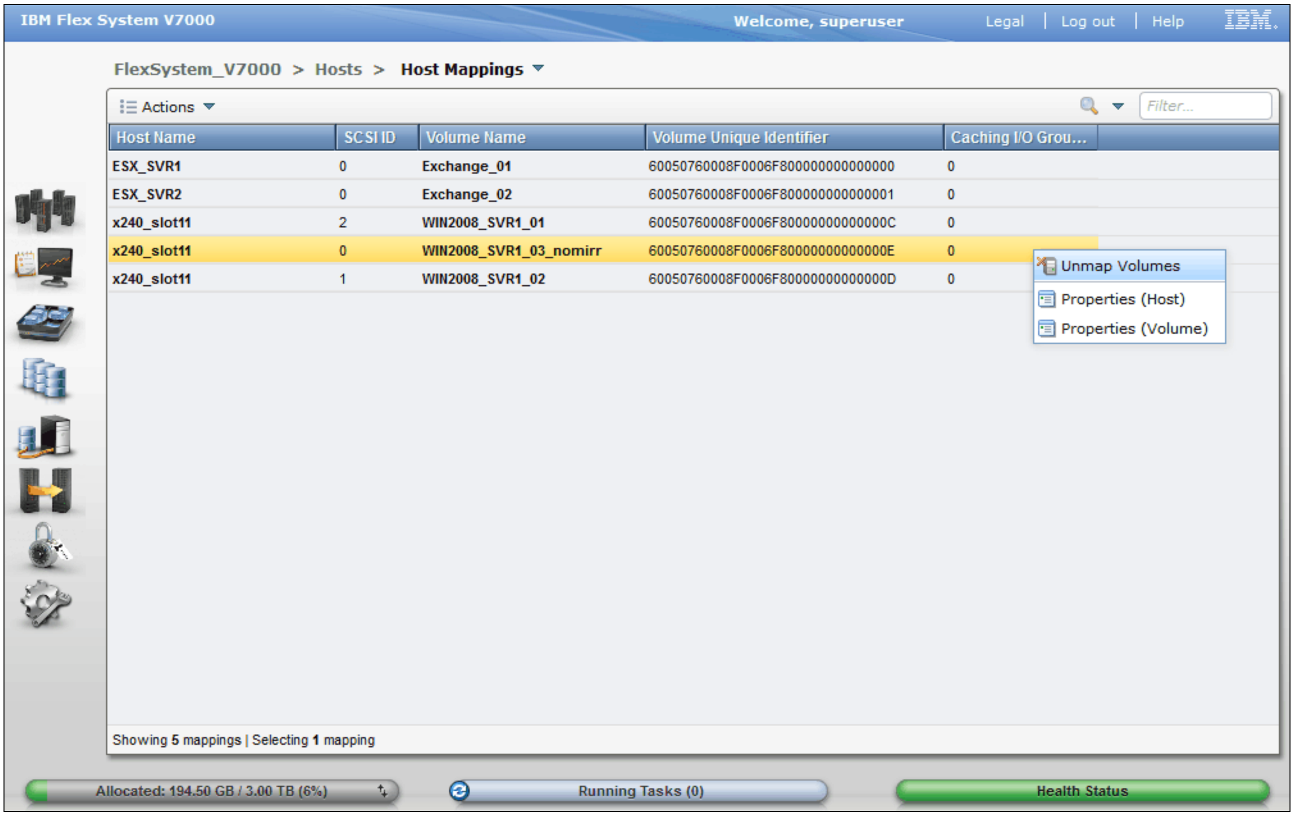


Figure 5-52 Actions for a single host from the Host Mapping window

By unmapping the volume as the example shown above, the volume will be made unavailable to the host.

5.6.4 Volumes by Host

The **Volumes by Host** menu is used to filter the view to show which hosts has which volumes mapped. Examples of how to use the **Volumes by Host** menu refer to Chapter 7, “Migration wizard” on page 265.

5.7 Copy Services menu

IBM Flex System V7000 Storage Node provide a number of different methods of coping and replicating data. FlashCopy is provided for instant copy of block volumes within the cluster. Remote copy is used to copy block volumes to another location on another cluster and this can be done synchronously (Metro Mirror) or asynschronously (Global Mirror).

In this section, we describe the Copy Services menu and its options. If you hover the cursor over the Copy Services function icon, the window shown in Figure 5-53 opens.

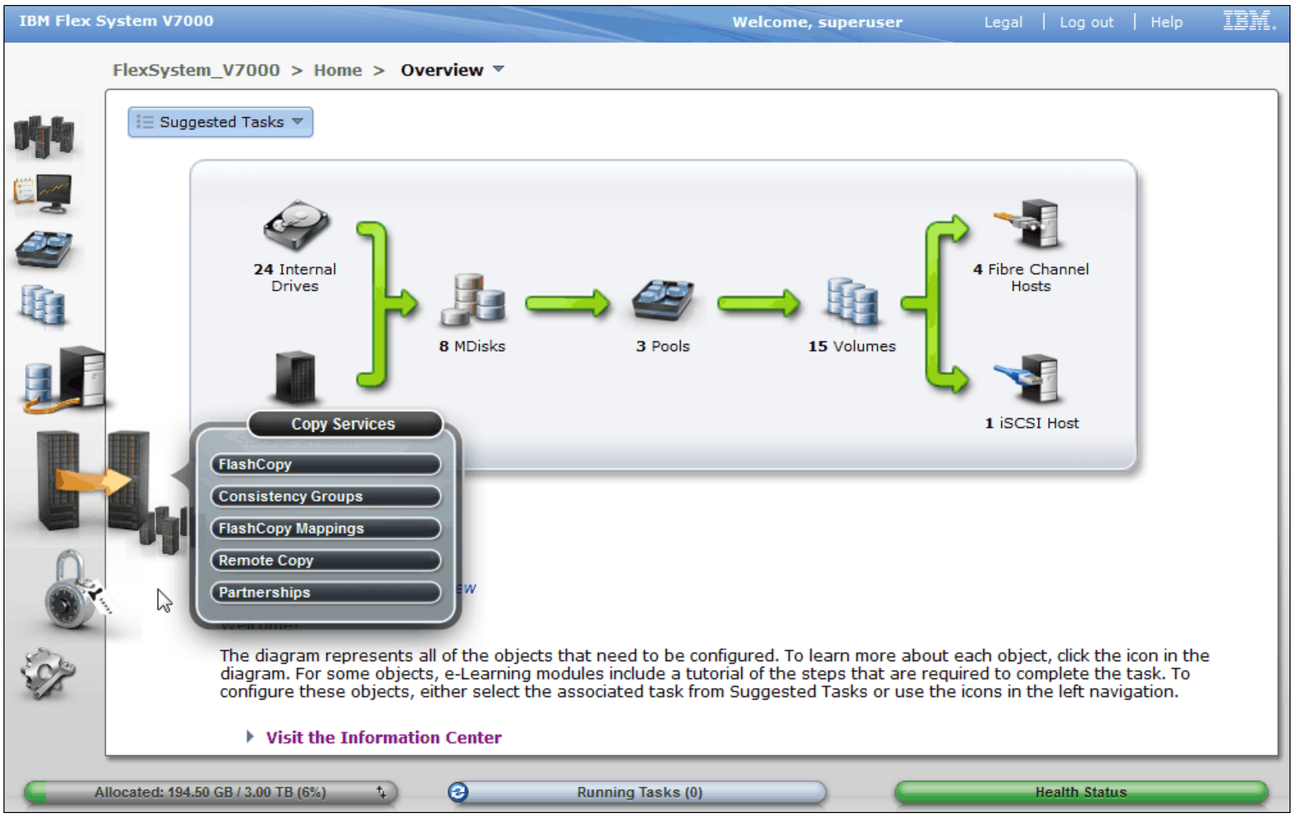


Figure 5-53 Navigate to the Copy Services menu

The actual features and functions of the Copy Services menu are discussed more detailed in Chapter 9, “IBM Flex System V7000 Storage Node Copy Services” on page 341.

5.7.1 FlashCopy menu

Clicking **FlashCopy** opens the window shown in Figure 5-54. This window shows the volumes that are available. If you right-click a volume, a list of operations opens. From here you can perform tasks such as initiate a new snapshot, clone or back up of a volume.

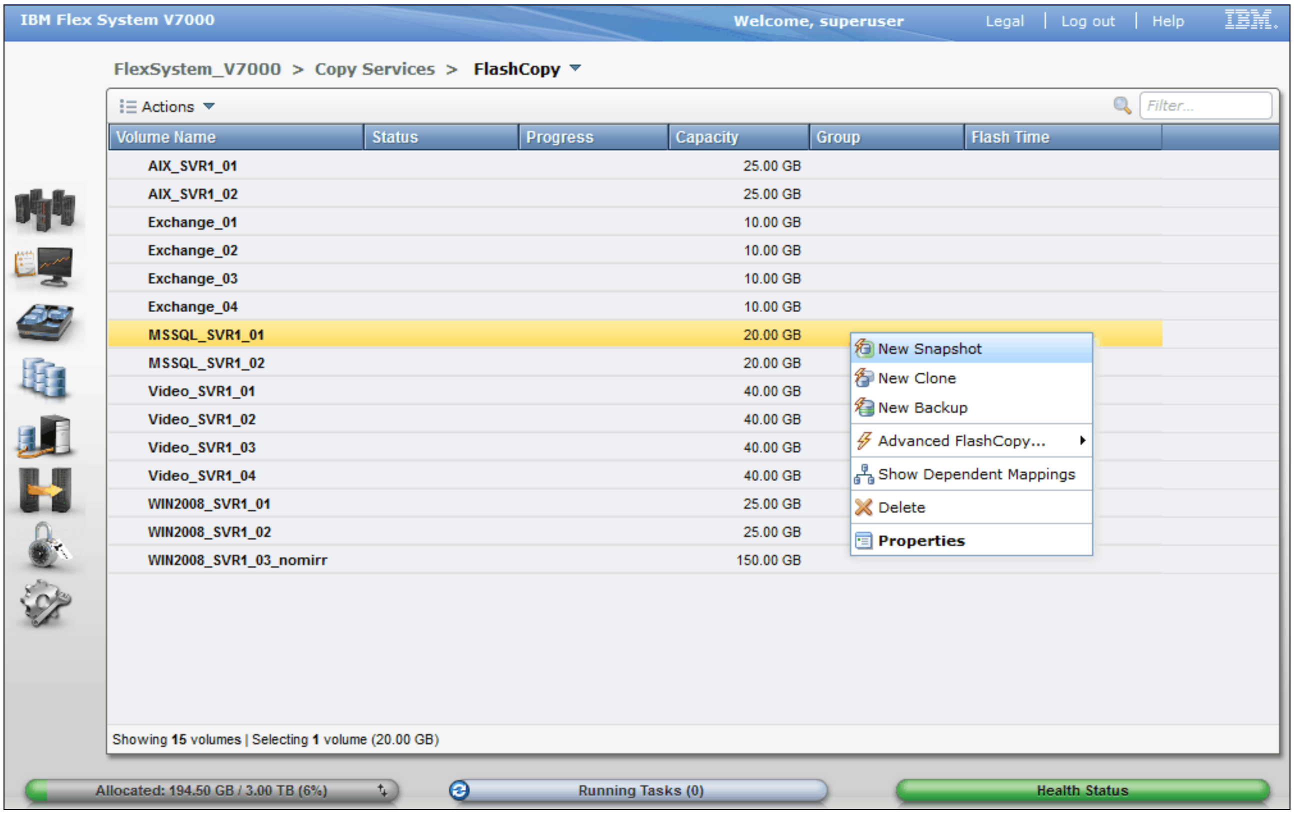


Figure 5-54 FlashCopy menu

By either double clicking the volume name or **Right click** → **Properties** opens the Volume Details window shown in Figure 5-55. From here, you can click the tabs at the top of the window to display additional information, such as the hosts that the volume or FlashCopy volume is mapped to and its dependant MDisks.

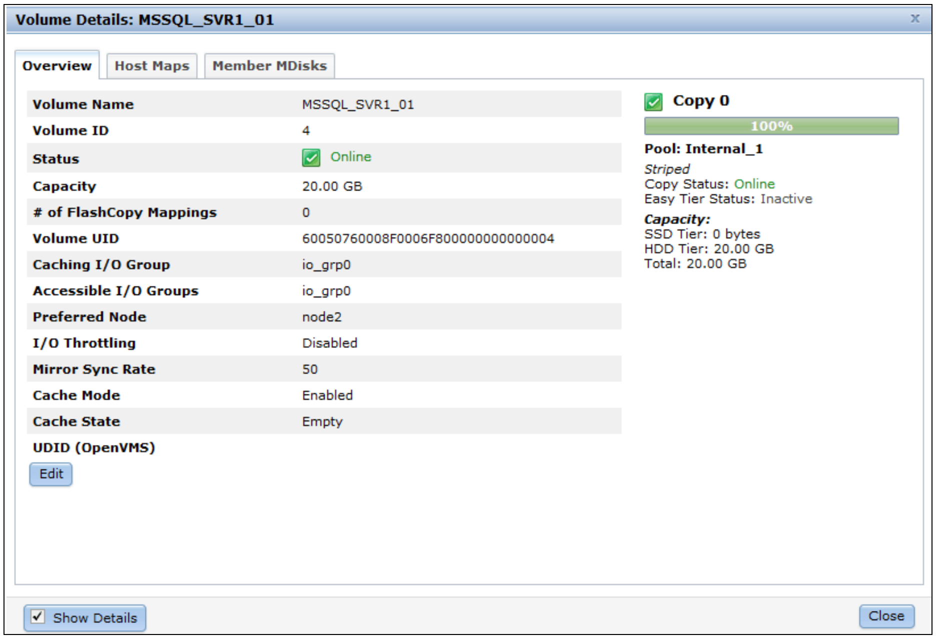


Figure 5-55 FlashCopy volume details with Show Details checked

5.7.2 FlashCopy Consistency Group menu

FlashCopy Consistency Groups are used to group multiple copy operations together that have a need to be controlled at the same time. In this way the group can be controlled by starting, stopping, and so on, with a single operation. Additionally, the function will ensure that when stopped for any reason, the IOs to all group members have all stopped at the same “point-in-time” in terms of the host writes to the primary volumes, ensuring time consistency across volumes.

Clicking **FlashCopy** → **Consistency Group** opens the window shown in Figure 5-56. This window enables FlashCopy relationships to be placed into a consistency group. You can also use start and stop commands against the FlashCopy consistency group from this window by right-clicking the relationship.

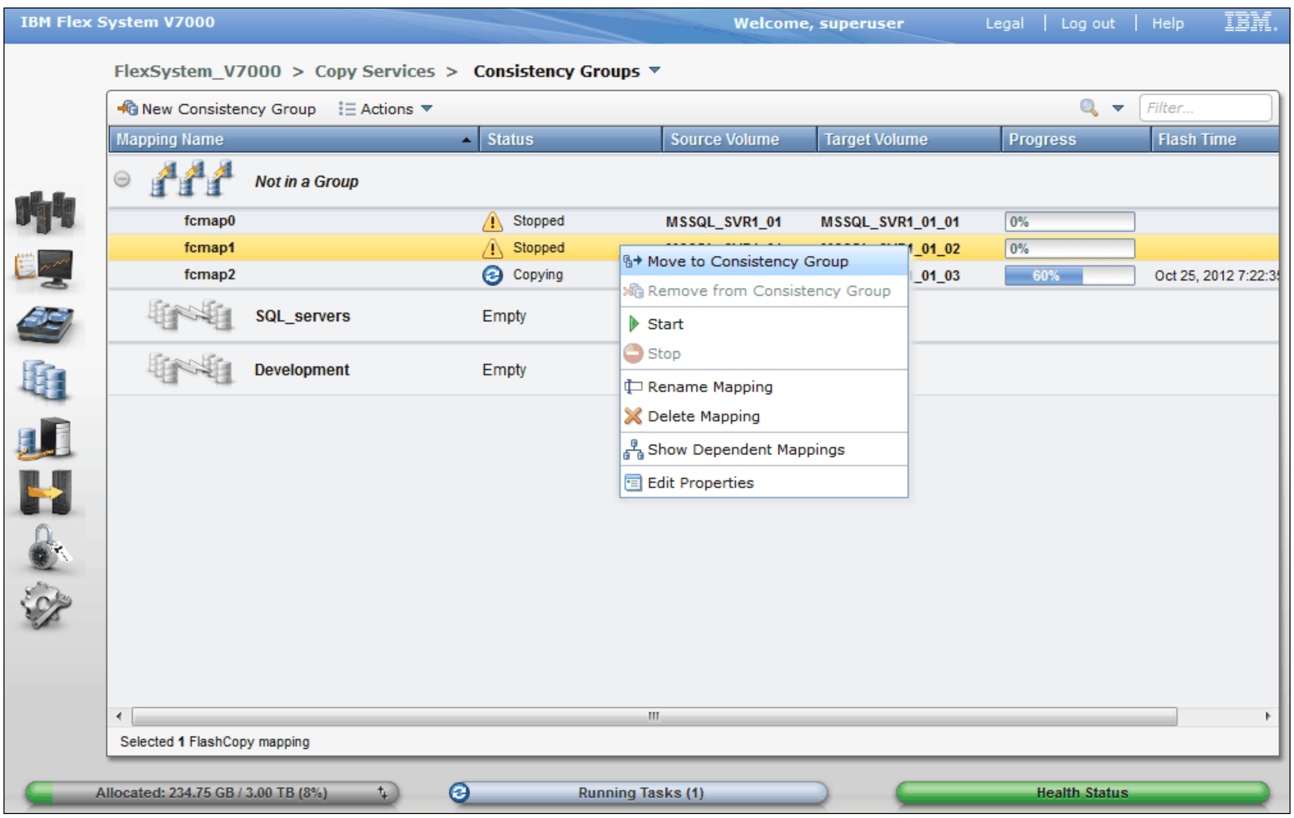


Figure 5-56 FlashCopy Consistency Group window

Consistency groups address the problem of applications having related data that spans multiple volumes. In this situation, FlashCopy operations must be performed in a way that preserves data integrity across the multiple volumes. One requirement for preserving the integrity of data being written is to ensure that dependent writes are run in the intended sequence of the application.

5.7.3 FlashCopy Mapping menu

The FlashCopy Mapping menu allows you to create and view the relationship (mapping) between the FlashCopy source and target volumes.

Clicking **FlashCopy Mapping** opens the window shown in Figure 5-57. From this window, you can Start, Stop, Delete, and Rename the FlashCopy mappings. There is also an option to move the relationship into a consistency group.

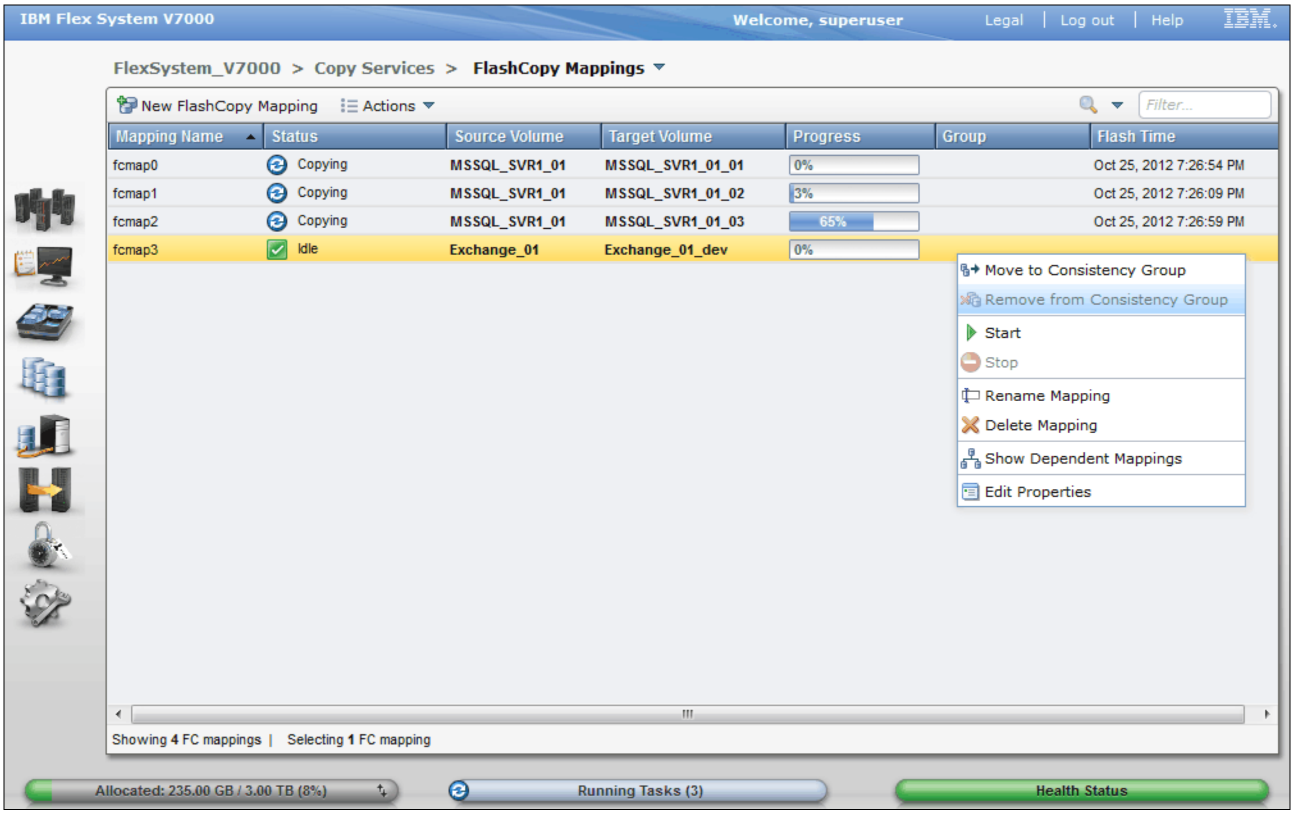


Figure 5-57 FlashCopy Mapping window

5.7.4 Remote Copy and the Partnerships

Remote Copy and the Partnerships menu are described in detail in 9.3, "Remote Copy" on page 380.

5.8 Access menu

There are a number of levels of user access to the IBM Flex System V7000 Storage Node cluster, which are managed through the access menu. The access levels are divided into groups each having a different level of access and authority. If desired, multiple users can be defined and their access assigned to suit the tasks they perform.

In this section, we describe the Access menu and its options. If you hover the cursor over the Access function icon, a menu opens (Figure 5-58).

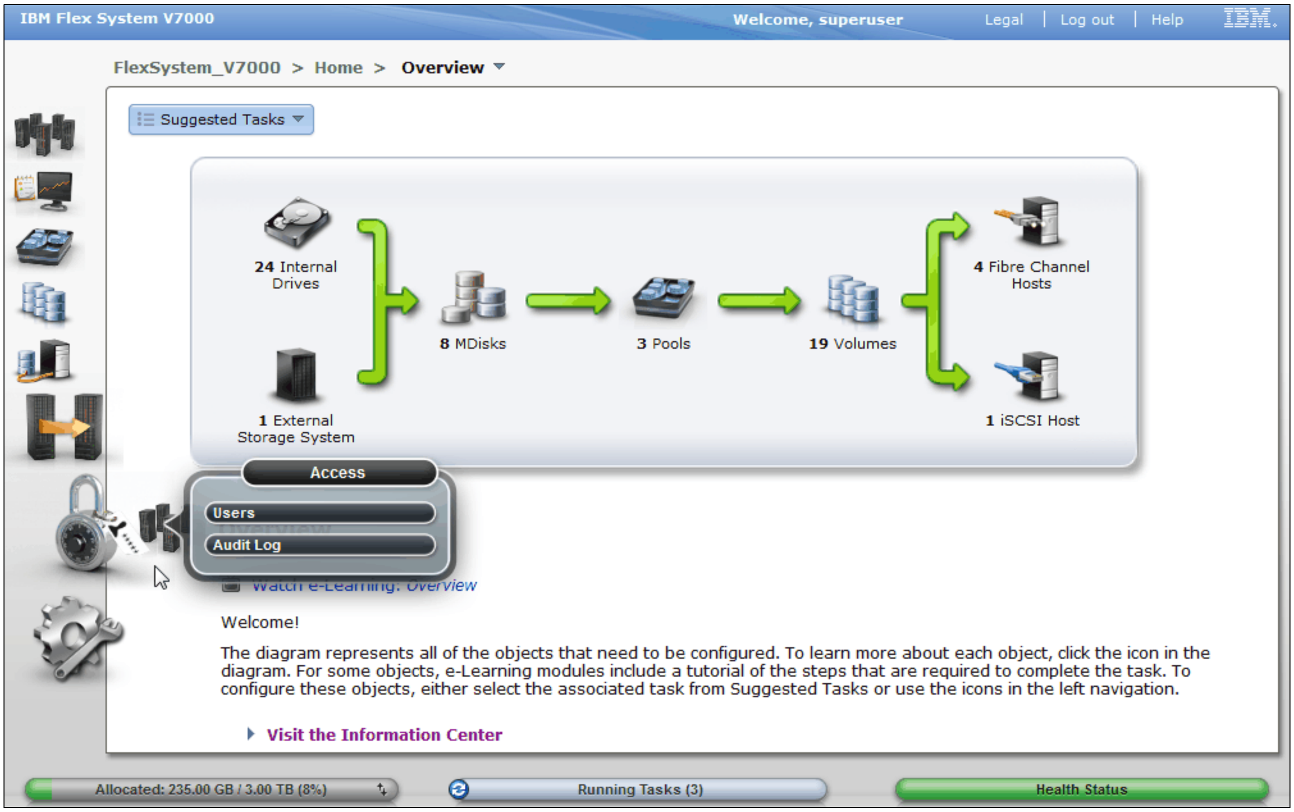


Figure 5-58 Navigate to the Access menu

The Access menu allows for user management and Audit log review.

User management includes creation of new users as well as maintaining roles and passwords for existing users. Also part of user management is configuration of SSH keys which provides secure access for users to the Command Line Interface (CLI)

The Audit log provides a list of all commands being executed on the system, and also contain information about which user ran the command.

5.8.1 Users menu

Figure 5-59 shows the Users window. This window enables you to create and delete new users, change and remove passwords, and add and remove SSH keys.

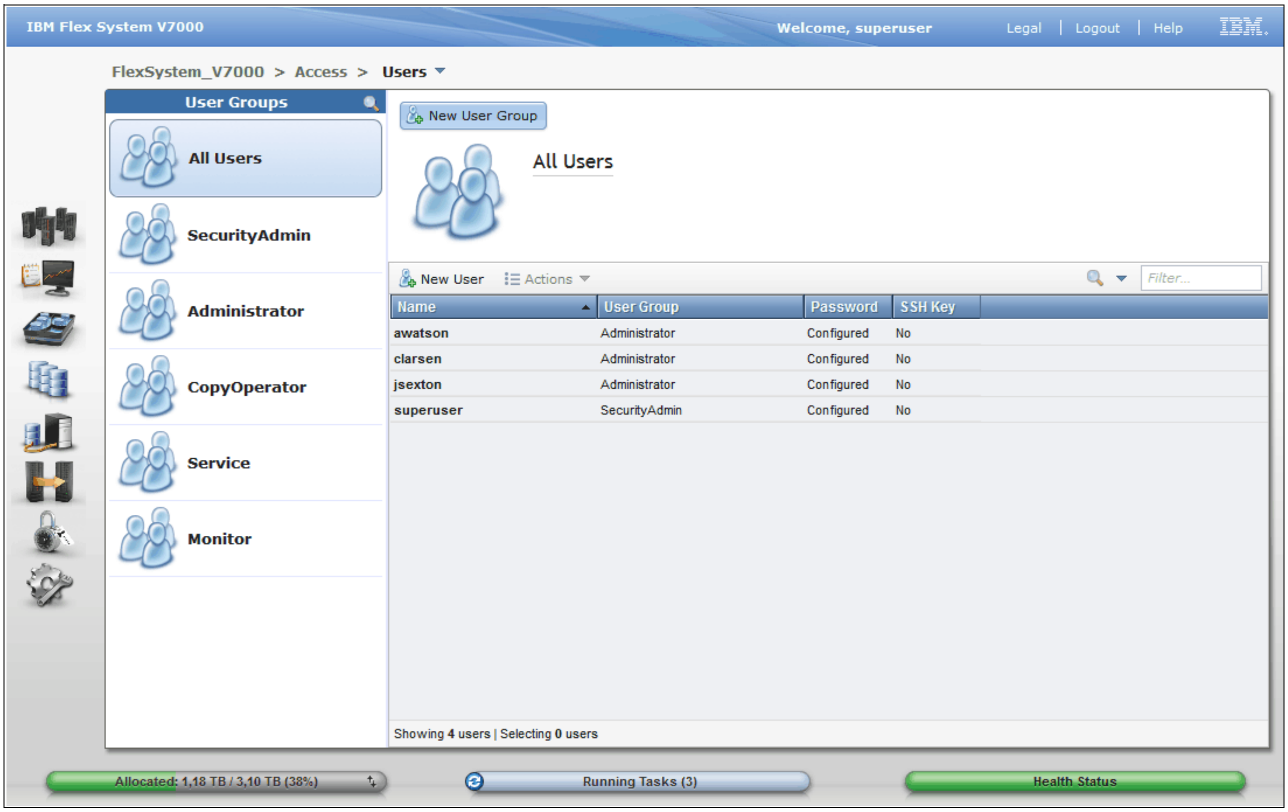


Figure 5-59 Users menu

Clicking **New User** opens the window shown in Figure 5-60. From here you can enter the name of the user and the password, and load the SSH key (if the SSH key has been generated). Starting with SAN Volume Controller and IBM Storwize V7000 V6.3 SSH keys are not required for CLI access, and you can choose either to use SSH or password for CLI authentication.

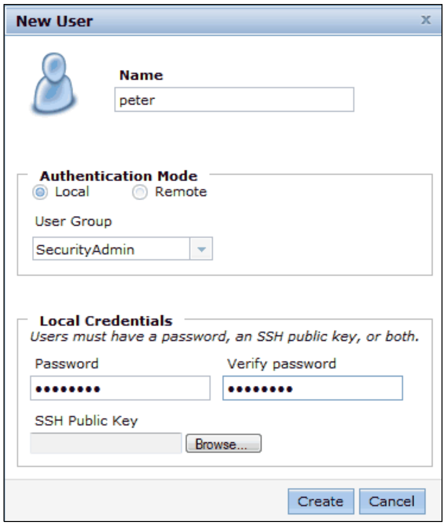


Figure 5-60 New User creation window

Configuring Command Line Interface access including how to configure SSH keys for secure access only is described in detail in Appendix A, “CLI Setup and Configuration” on page 573.

5.8.2 Audit Log menu

All commands issued on the cluster are logged in the Audit log. Note that even if initiated from the GUI, most actions cause a CLI command to be run, so this will also be logged.

Clicking **Audit Log** opens the window shown in Figure 5-61. The cluster maintains an audit log of successfully executed commands, indicating which users performed particular actions at certain times.

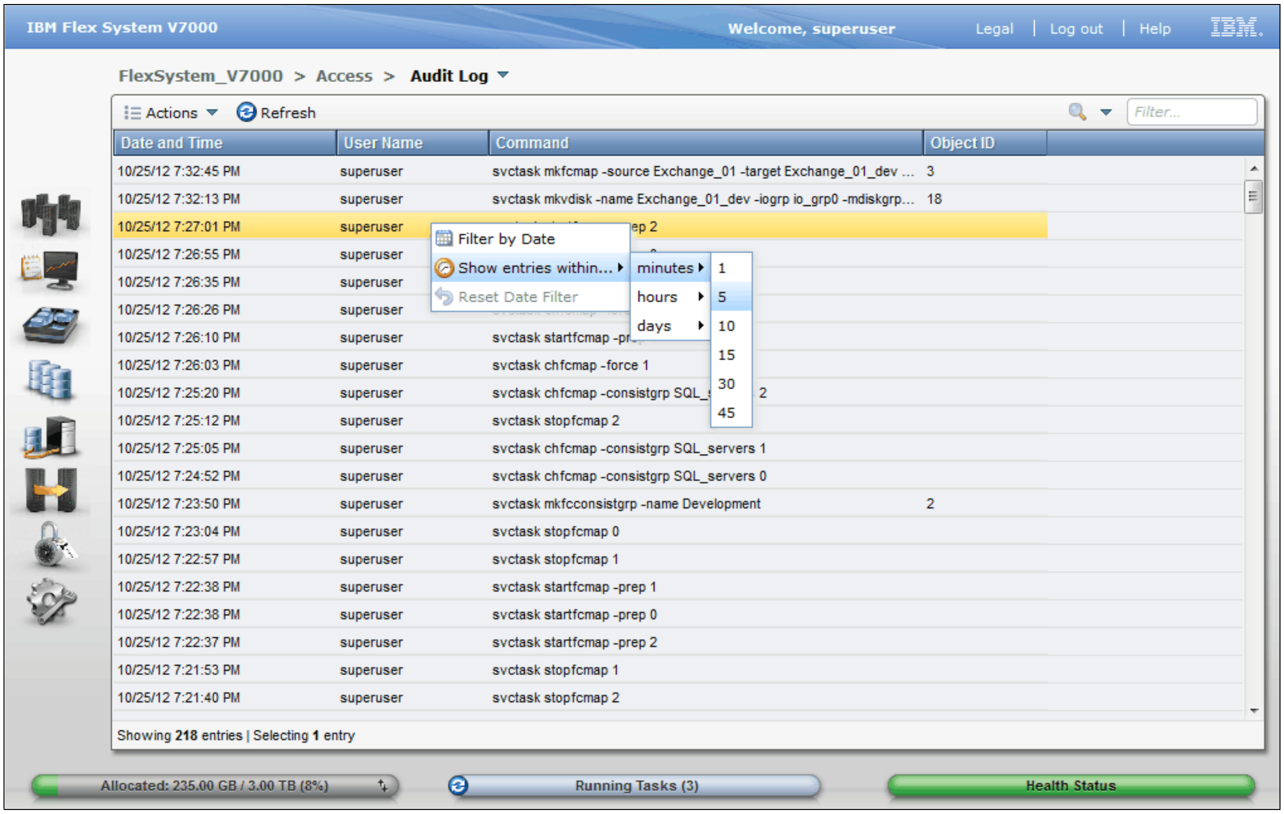


Figure 5-61 Viewing Audit Log menu using filtering

5.9 Settings menu

In the following section, we describe the Settings menu and its options. If you hover the cursor over the Settings function icon, the Settings menus opens (Figure 5-62).

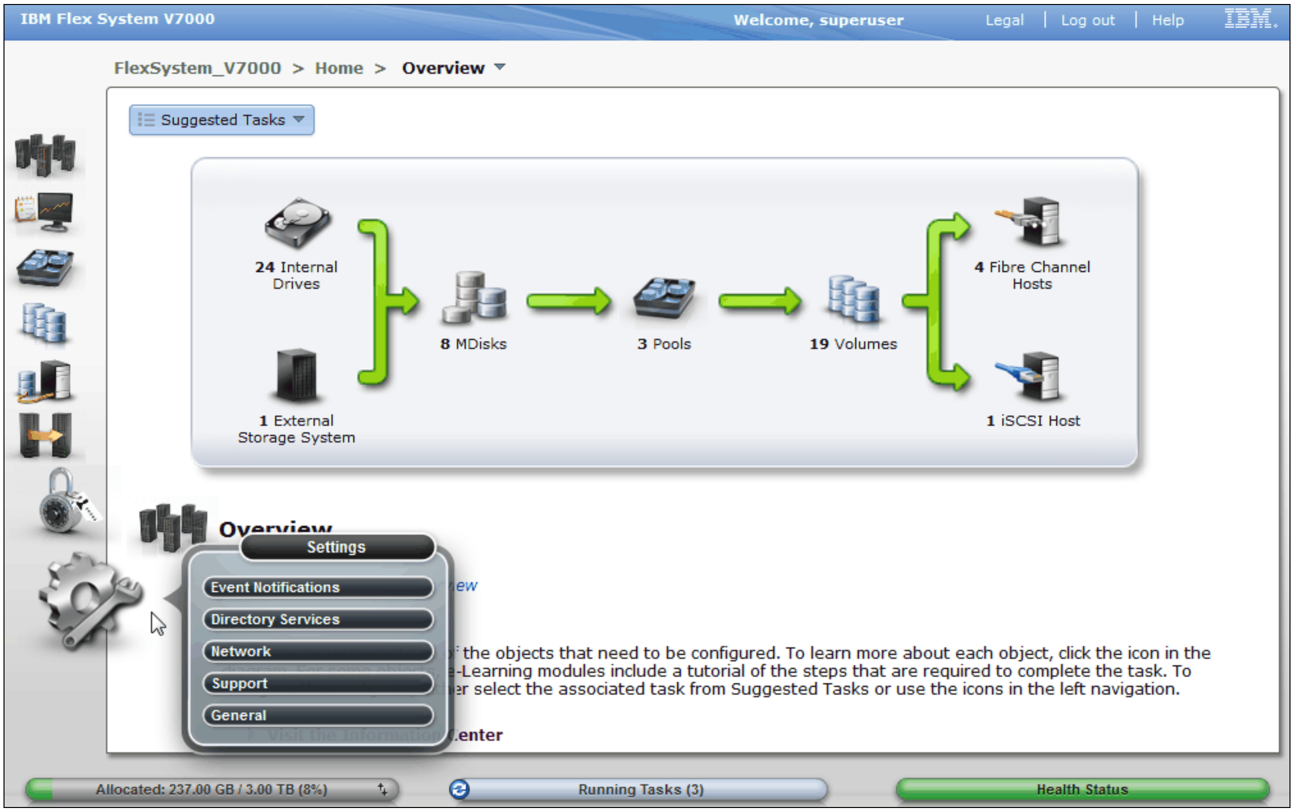


Figure 5-62 Navigate to the Settings menu

5.9.1 Event Notification menu

This option is used to configure the alerting and logging. Here we define the e-mail and SNMP servers and the levels of alerting as desired.

Clicking **Event Notification** opens the window shown in Figure 5-63. From this window, you can configure email alerts (which is included in the Call Home function) and SNMP monitoring, and define syslog servers and the message types.

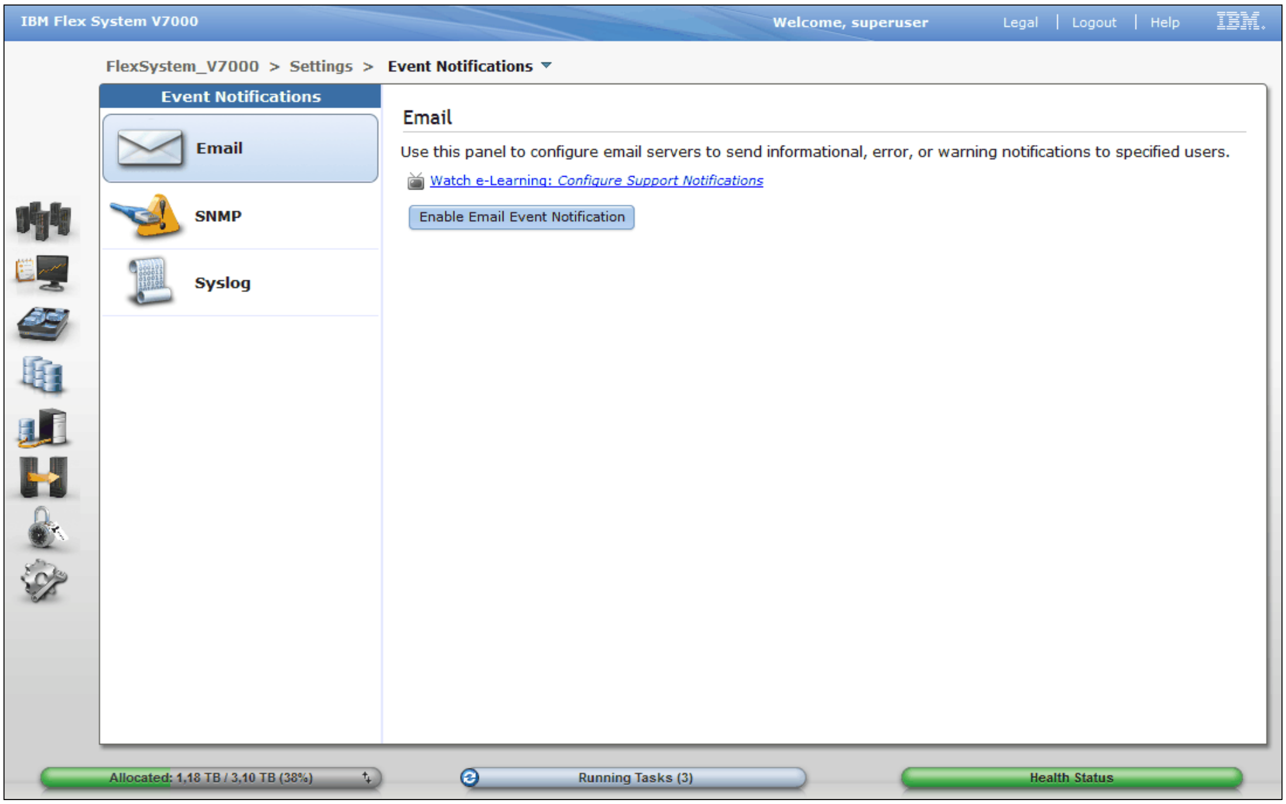


Figure 5-63 Event notification window

For more information about call home refer to Chapter 13, “Maintenance and Troubleshooting” on page 523.

5.9.2 Directory Services

When a IBM Flex System V7000 Storage Node clustered system is created, the authentication settings are automatically provisioned from the settings provided by the Chassis Management Module (CMM) or Flex System Manager (FSM). Users can log in to the system using the user accounts that are defined on the LDAP servers which are configured on the CMM or FSM.

For users of the Flex System V7000 Storage Node clustered system you can configure authentication and authorization using the Command Line Interface.

You can create two types of users who can access the system. These types are based on how the users are authenticated to the system. Local users must provide either a password, a Secure Shell (SSH) key, or both. Local users are authenticated through the authentication methods that are located on the Flex System V7000 Storage Node system. If the local user

needs access to the management GUI, a password is needed for the user. If the user requires access to the command-line interface (CLI) through SSH, either a password or a valid SSH key file is necessary. Local users must be part of a user group that is defined on the system. User groups define roles that authorize the users within that group to a specific set of operations on the system.

A remote user is authenticated on a remote service with Lightweight Directory Access Protocol (LDAP) as configured in the Chassis Management Module or Flex System Manager settings. Remote users have their groups or Supervisor role defined by the remote authentication service.

A remote user who needs access to the command line using SSH keys must be configured on the Flex System V7000. Other remote users do not need to be configured on the V7000, they only need to be defined on the LDAP server.

To manage users and user groups on the system using the management GUI, select User **Management** → **Users**.

The settings for Lightweight Directory Access Protocol are automatically configured by the Chassis Management Module. To test the connection select **Settings** → **Directory Services**.

For more information of how to configure Remote Authentication and authorization for users of the IBM Flex System V7000 Storage Node refer to the *User Authentication Configuration* section of IBM Infocenter at
http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/svc_webconfiguringauthusers_224e4a.html

For more information on how to Implement Systems Management of IBM PureFlex Systems refer to*Implementing Systems Management of IBM PureFlex System*, SG24-8060.

5.9.3 Network menu

The network setup for all the interfaces in the cluster is configured here.

Clicking **Network** opens the window shown in Figure 5-64. From here you can update the network configuration, set up iSCSI definitions, and view information about the Fibre Channel connections.

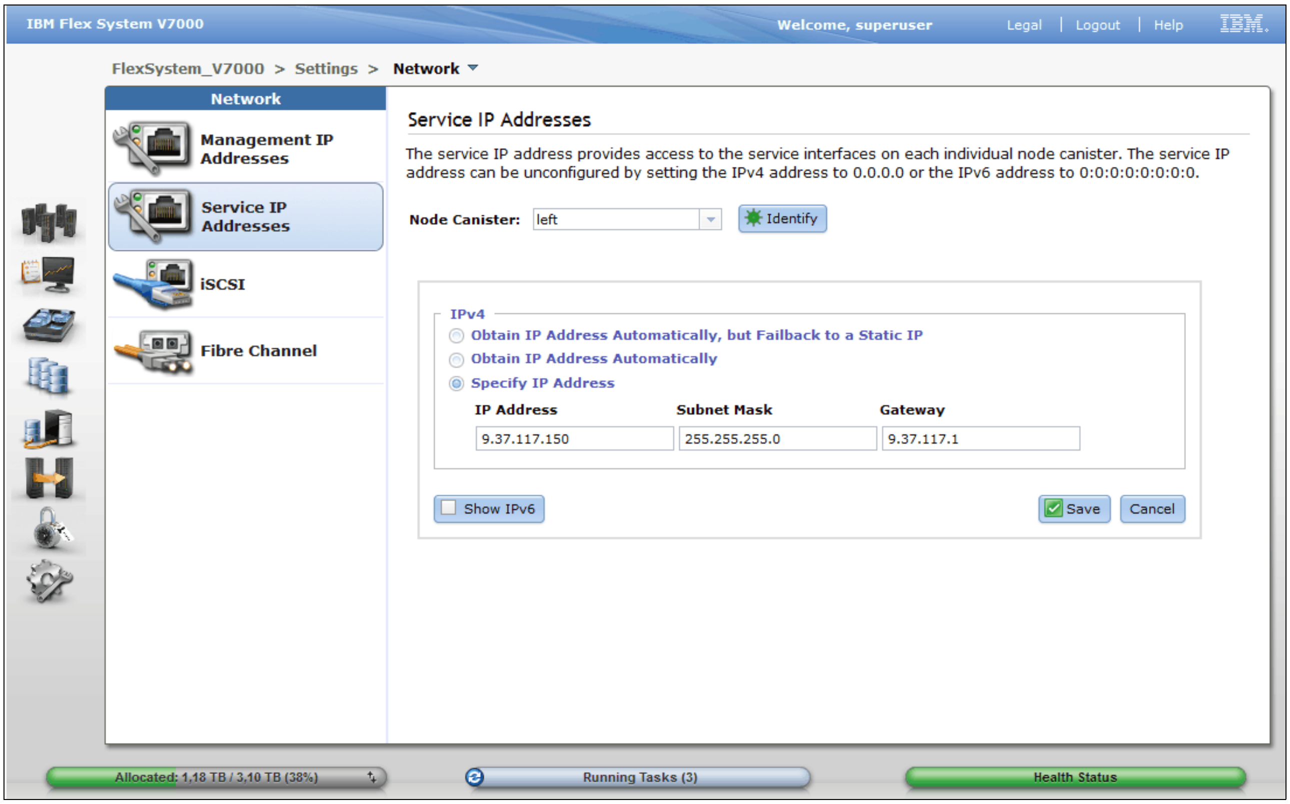


Figure 5-64 Network window

Service IP Addresses

The Service IP's are IP addresses of each of the two IBM Flex System V7000 Storage Node controllers called canisters. These controllers have their own IP addresses where several support actions can be performed. Some of these are listed here:

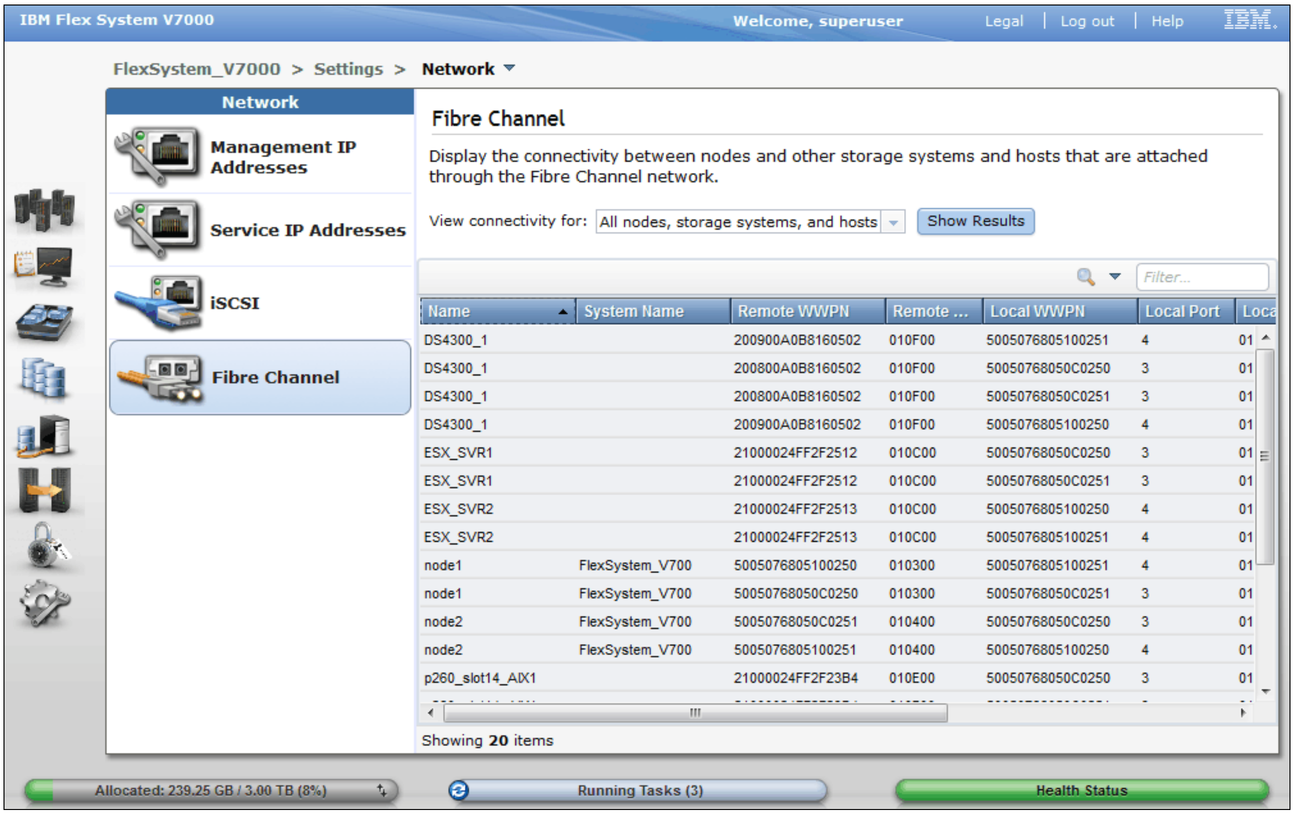
- Place canister in Service State
- Power cycle canister
- Clear all system configuration data
- Create new cluster
- Recover a failed system - action ONLY performed by IBM Support
- Upgrade firmware manually with controllers offline
- Extract system event logs

> **Note:** The Service IP addresses are normally not used by the IBM Flex System V7000 Storage Node administrator. They are used ONLY as part of the initial installation procedure or when IBM Support performs certain service actions.

Fibre Channel

Fibre Channel is used to display the FC-paths and their status. This is the Fibre Channel paths over the SAN and only relates to the storage nodes.

When you click **Fibre Channel** (Figure 5-65), some useful information is displayed about Fibre Channel connectivity. In this example, we review connectivity for all nodes, storage systems and hosts. Views can be filtered to display details for all devices, systems, nodes, storage systems, and hosts.



Name	System Name	Remote WWPN	Remote ...	Local WWPN	Local Port	Loca
DS4300_1		200900A0B8160502	010F00	5005076805100251	4	01
DS4300_1		200800A0B8160502	010F00	50050768050C0250	3	01
DS4300_1		200800A0B8160502	010F00	50050768050C0251	3	01
DS4300_1		200900A0B8160502	010F00	5005076805100250	4	01
ESX_SVR1		21000024FF2F2512	010C00	50050768050C0250	3	01
ESX_SVR1		21000024FF2F2512	010C00	50050768050C0251	3	01
ESX_SVR2		21000024FF2F2513	010C00	5005076805100250	4	01
ESX_SVR2		21000024FF2F2513	010C00	5005076805100251	4	01
node1	FlexSystem_V700	5005076805100250	010300	5005076805100251	4	01
node1	FlexSystem_V700	50050768050C0250	010300	50050768050C0251	3	01
node2	FlexSystem_V700	50050768050C0251	010400	50050768050C0250	3	01
node2	FlexSystem_V700	5005076805100251	010400	5005076805100250	4	01
p260_slot14_AIX1		21000024FF2F23B4	010E00	50050768050C0250	3	01

Figure 5-65 Fibre Channel window

5.9.4 Support menu

The Support menu is mainly used when log files are requested by IBM Support. IBM Support often requests log files when a support case is opened by IBM Flex System V7000 Storage Node administrators. The overall process is that the system administrator download the requested support package from the system and then upload it to IBM Support where after the data can be analyzed.

Clicking **Support** opens the window shown in Figure 5-66 on page 235. From this window, click **Show full log listing** to show all the log files.

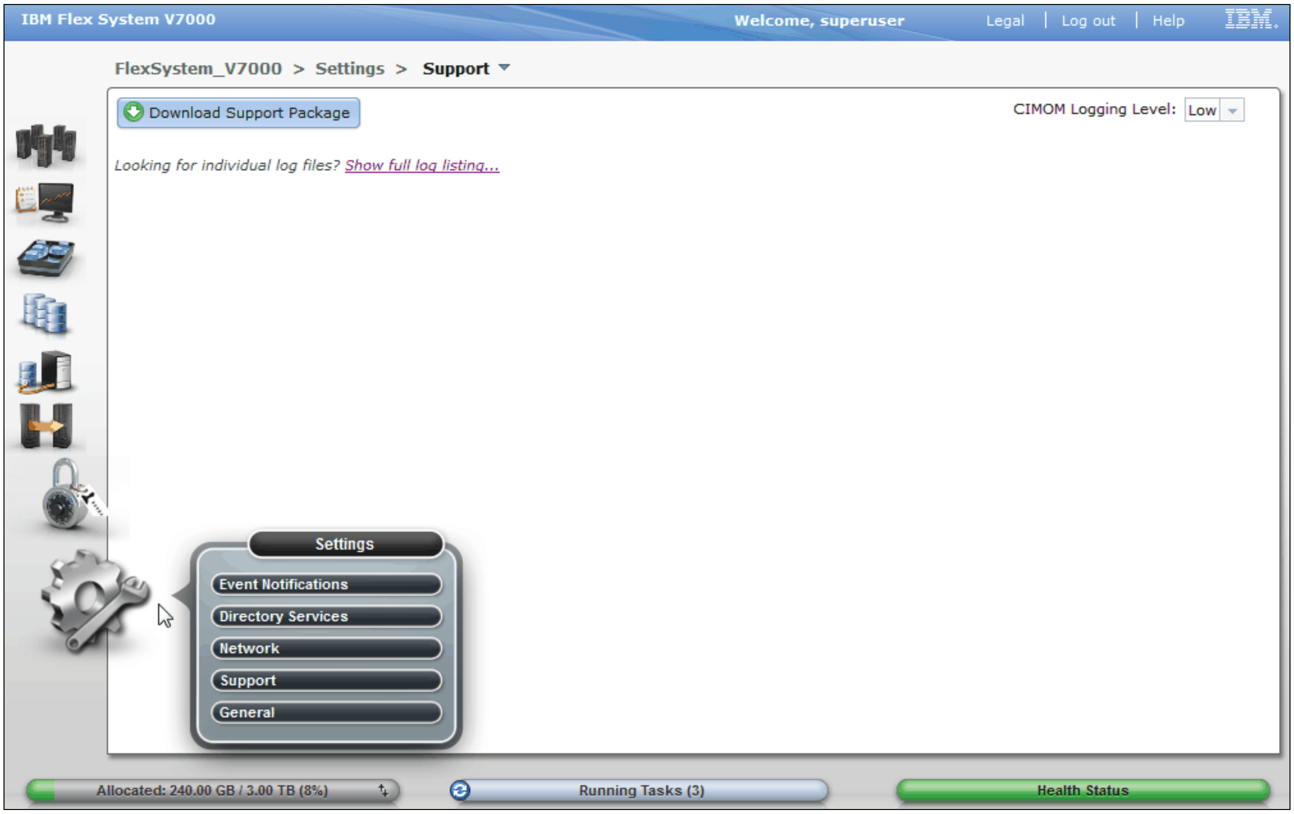


Figure 5-66 Support menu

You can download any of the various log files or delete them by selecting a single item (Figure 5-67) and click either the **Download** or **Delete** options under the Actions button.

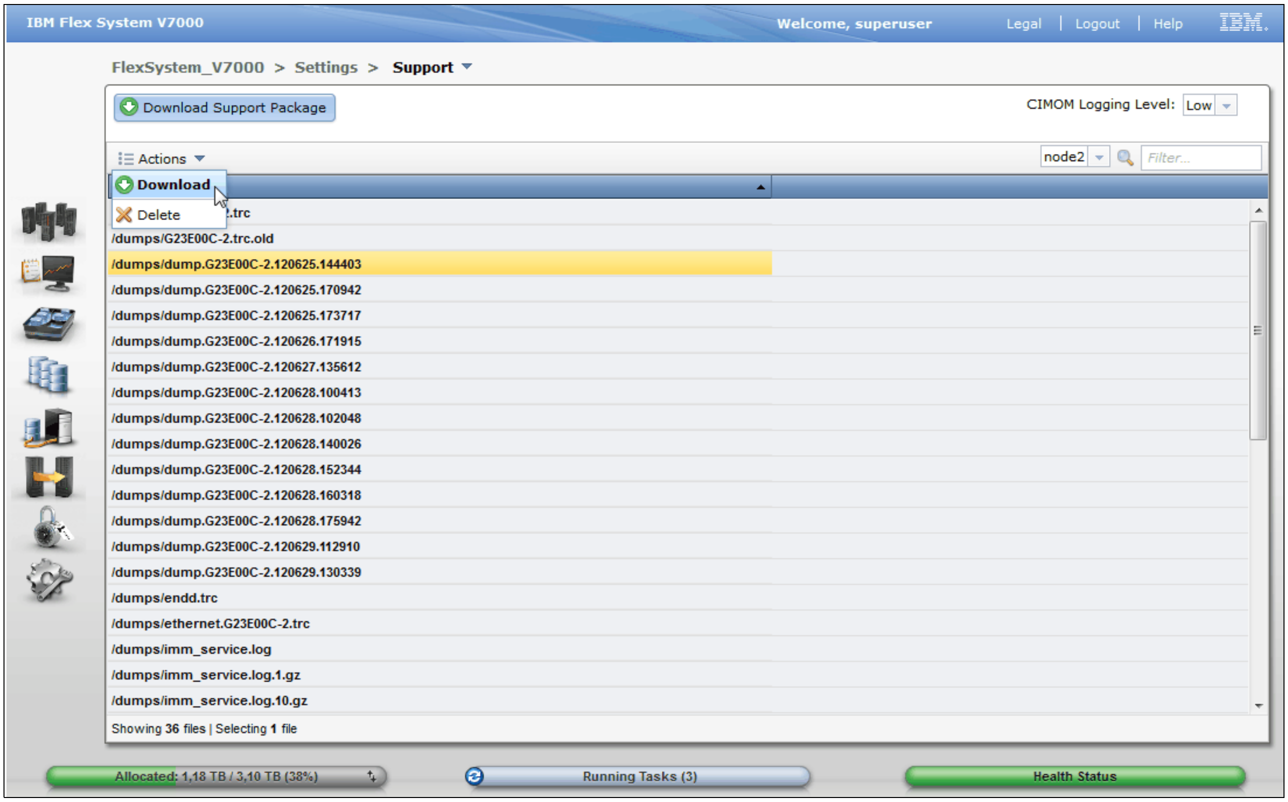


Figure 5-67 Download / Delete options of the Actions button

Delete option: When the Delete option is not available, the file cannot be deleted because it is a file being used by the system.

At the upper right of the window, there is a Node option that can be used to show node canister 1 or 2 log files (Figure 5-68).

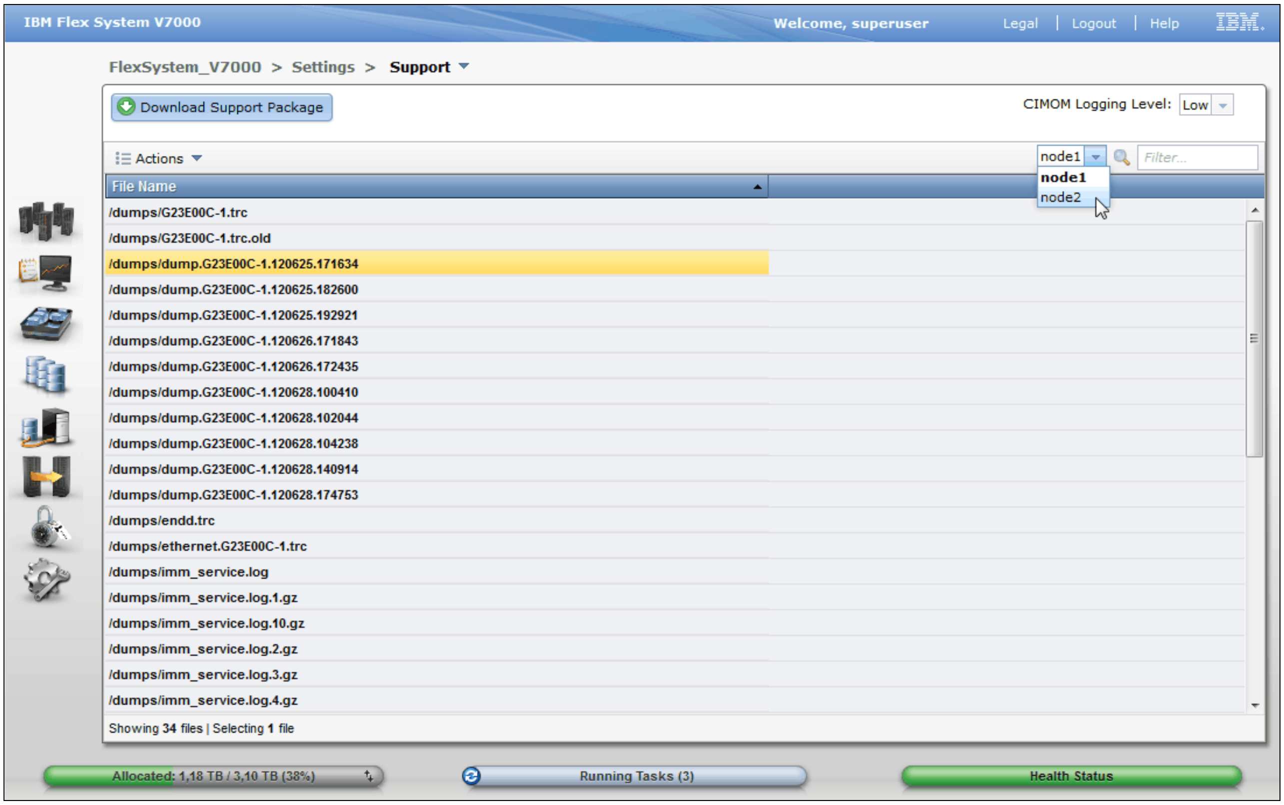


Figure 5-68 Change the log listing of the Nodes canister

Clicking **Download Support Package** opens the window shown in Figure 5-69. This window provides a number of different options that can be used to collect logs and save state information from the cluster.

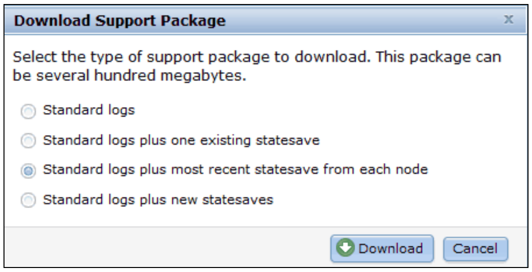


Figure 5-69 Download a support package

Clicking **Download** generates the support package shown in Figure 5-70.

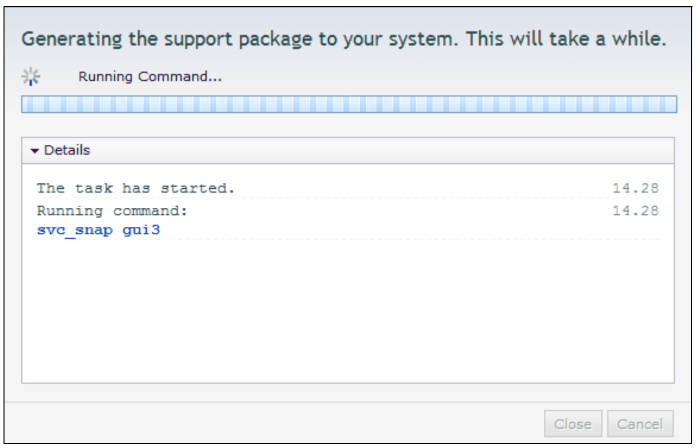


Figure 5-70 Generate a Support Package

Click **Save File** and then **Save** to save a copy of the package (Figure 5-71 on page 239).

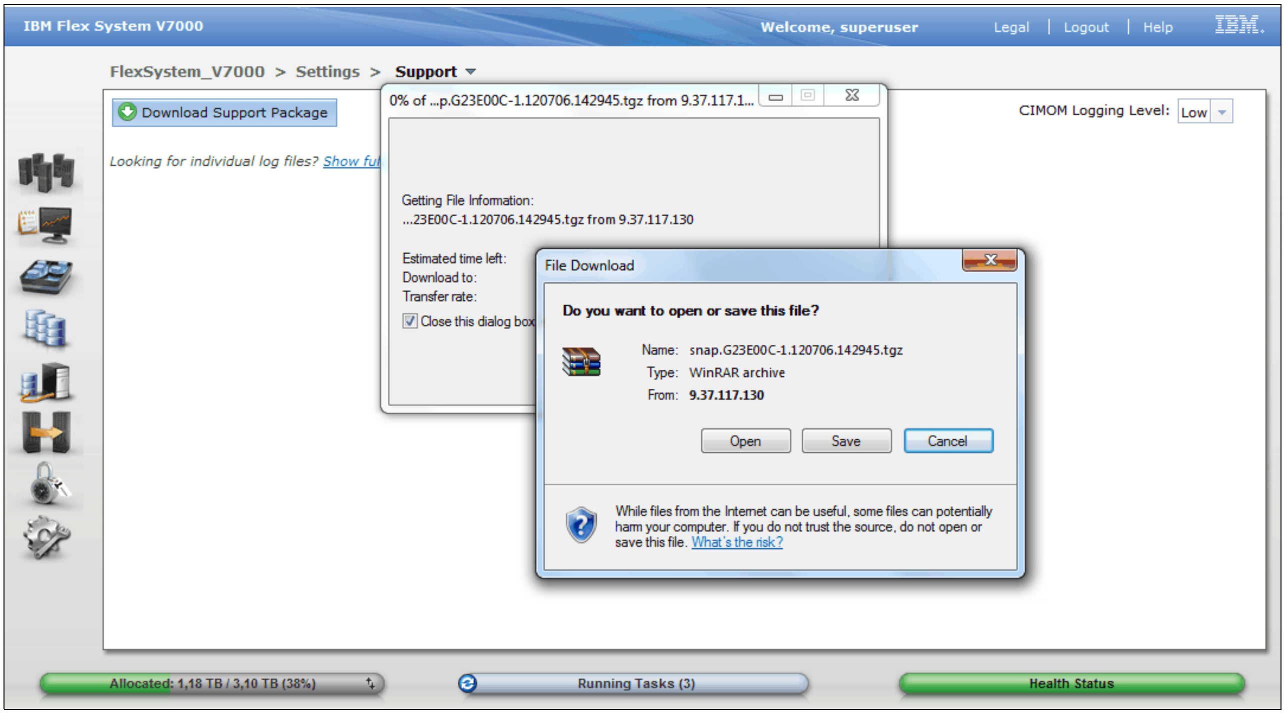


Figure 5-71 Save the Support Package

The downloaded support package is now ready for analysis which is part of how IBM Support deals with troubleshooting on the IBM Flex System V7000 Storage Node and similar IBM storage virtualization products.

5.9.5 General menu

In this option we can set the time and date for the cluster, enter licensing details if needed and perform software upgrades for the cluster.

Clicking the **General** option opens the window shown in Figure 5-72. This window provides options to set the date and time, update the software licensing levels, upgrade the firmware, and set GUI preferences.

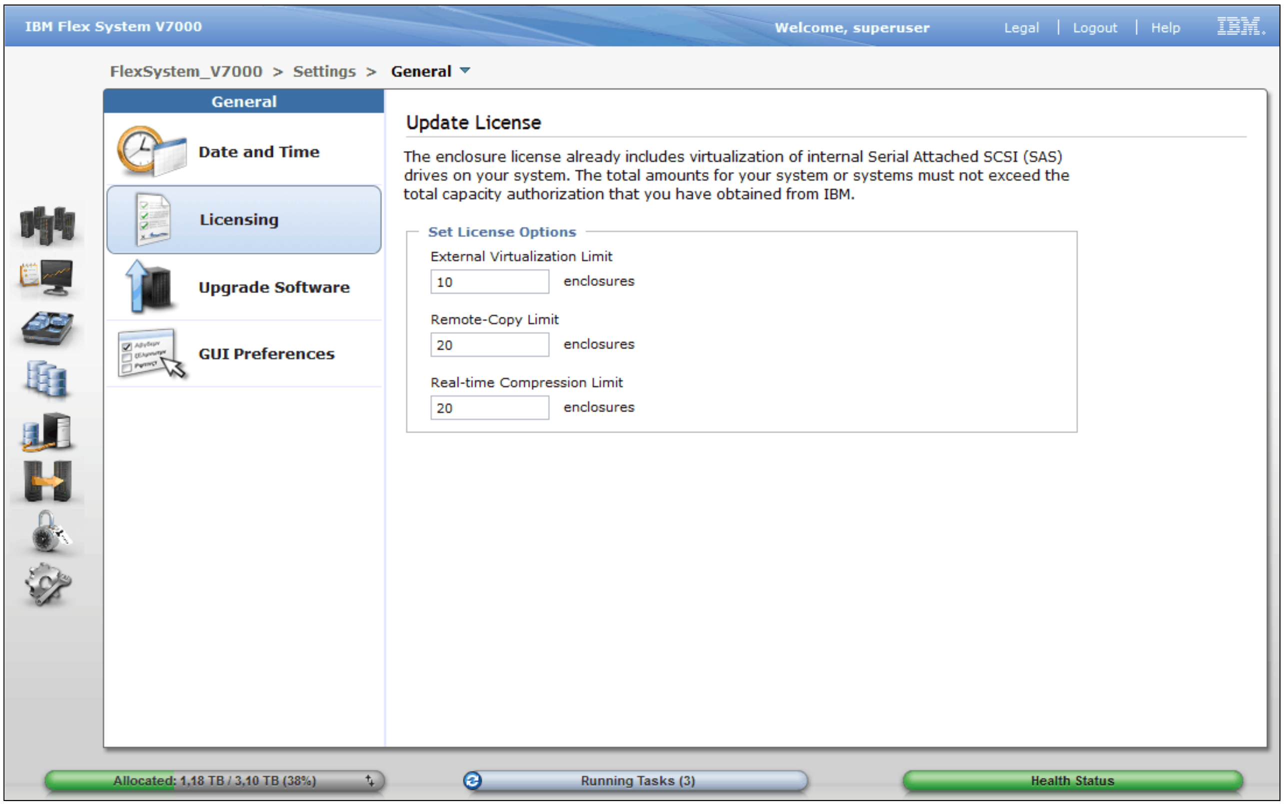


Figure 5-72 General menu showing licensing

6

Basic Volume and Host Configuration

This chapter describes how to use the IBM Flex System V7000 Storage Node to create a volume and map a volume to a host. A volume is a logical disk on the IBM Flex System V7000 Storage Node that is provisioned out of a storage pool and is recognized by a host with an identifier UID field and a parameter list.

The first part of this chapter describes how to create the different type of volumes that can be provided by IBM Flex System V7000 Storage Node.

The second part of this chapter covers how to map volumes to the attached hosts and finally the last part covers how to discover those volumes from the hosts.

Migrating volumes is covered in Chapter 7, “Migration wizard” on page 265.

Advanced host and volume administration, such as creating volume copies, flashcopies and so on, is described in Chapter 9, “IBM Flex System V7000 Storage Node Copy Services” on page 341.

6.1 Storage provisioning from IBM Flex System V7000 Storage Node

This section follows the setup process and shows how to create volumes and make them accessible from the host. The following steps are required to complete the basic setup of your environment:

1. Create new volumes - starting at 6.1.1, “Creating a generic volume” on page 244
2. Create New Host - 6.2, “Creating a new host” on page 256
3. Map volumes to the host 6.3, “Mapping a volume to the host” on page 261

Discovery of the volumes from the host and multipath settings are discovered in Chapter 12, “Host Configuration” on page 473.

Open the All Volumes window of the IBM Flex System V7000 Storage Node GUI (Figure 6-1) to start the process of creating new volumes.

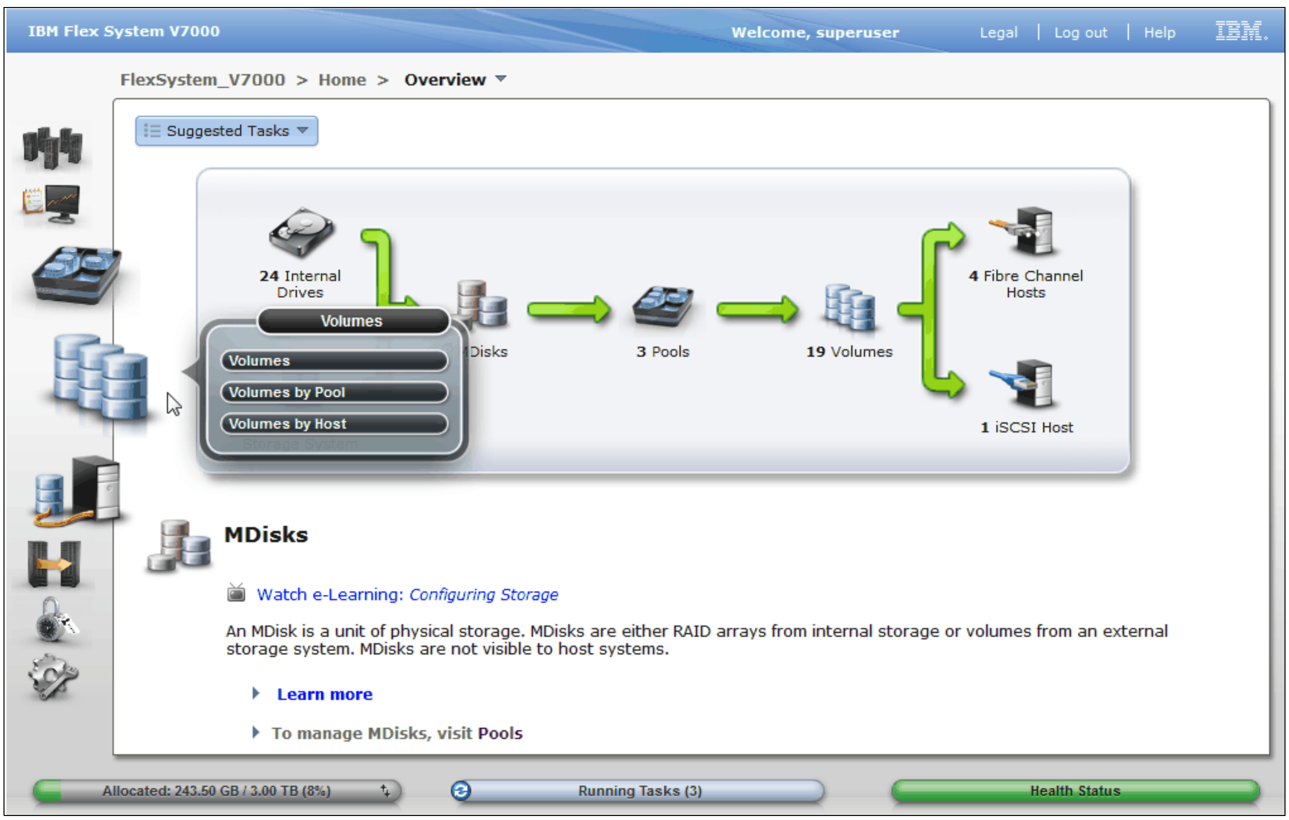


Figure 6-1 Home window

The All Volumes window opens (Figure 6-2). If this the first time volumes are configured on a newly initialized system, a pop up message appears that no volumes are defined yet.

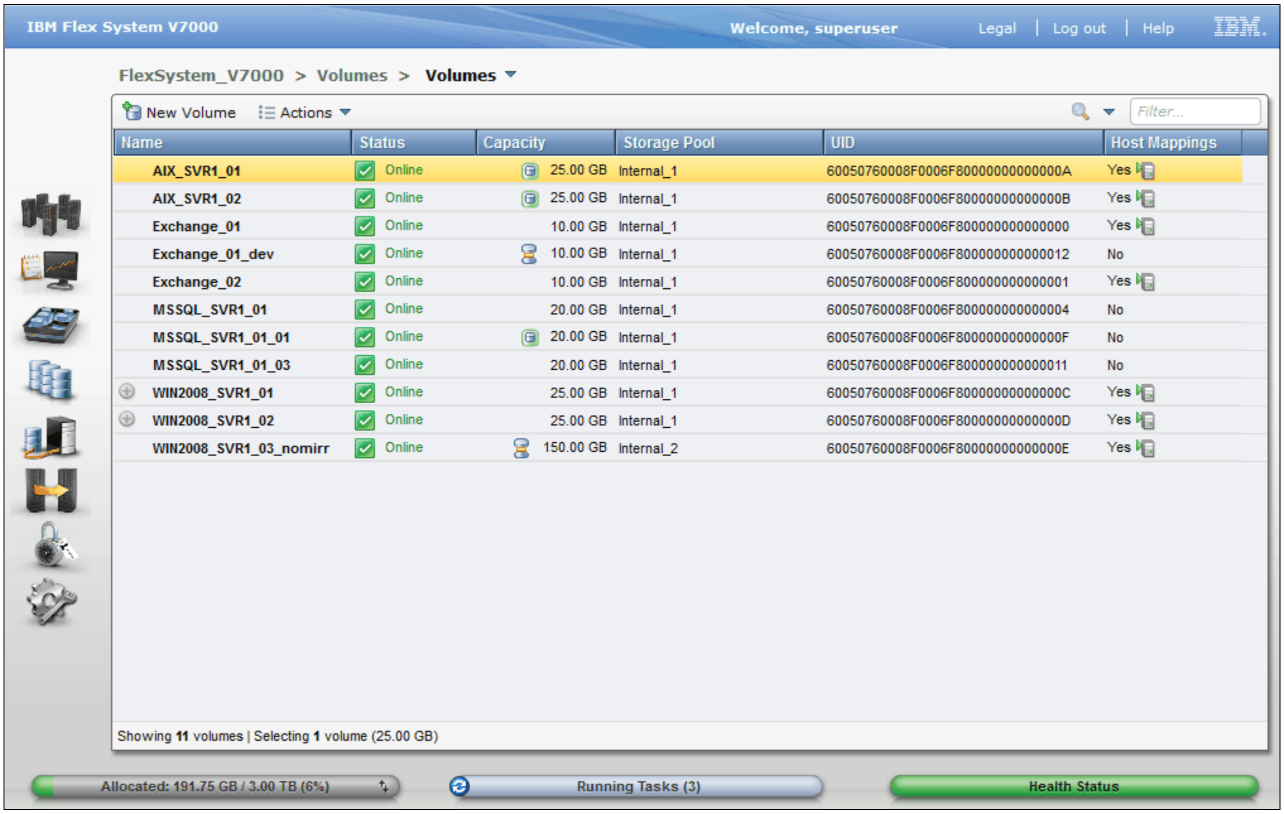


Figure 6-2 All Volumes window

We have some volumes already and are creating additional volumes in the following sections.

By default, all volumes that you create are striped across all available MDisks in one storage pool. The GUI for the IBM Flex System V7000 Storage Node provides the following preset selections for the user:

- Generic: A striped volume that is fully provisioned, as described in 6.1.1, “Creating a generic volume” on page 244.
- Thin-provisioned: A striped volume that is space efficient. There are choices under the Advanced button to help determine how much space is fully allocated initially and how large the volume is able to grow, as described in 6.1.2, “Creating a thin-provisioned volume” on page 245.
- Mirror: The striped volume consists of two striped copies and is synchronized to protect against loss of data if the underlying storage pool of one copy is lost, as described in 6.1.3, “Creating a mirrored volume” on page 248.
- Thin-mirror: Two synchronized copies. Both are thin provisioned, as described in 6.1.4, “Creating a thin-mirror volume” on page 251.
- Compressed: A volume using embedded Real-time Compression technology for reducing the size of the volume, as described in “Creating a compressed volume” on page 254

Select which volume you want to create and go to the relevant section.

6.1.1 Creating a generic volume

The most commonly used type of volume is the generic volume type. This type of volume is fully provisioned, with the entire size of the volume dedicated to the volume. The host and the IBM Flex System V7000 Storage Node see the fully allocated space without a mirror.

1. Click **New Volume** and select **Generic** as shown in Figure 6-3.

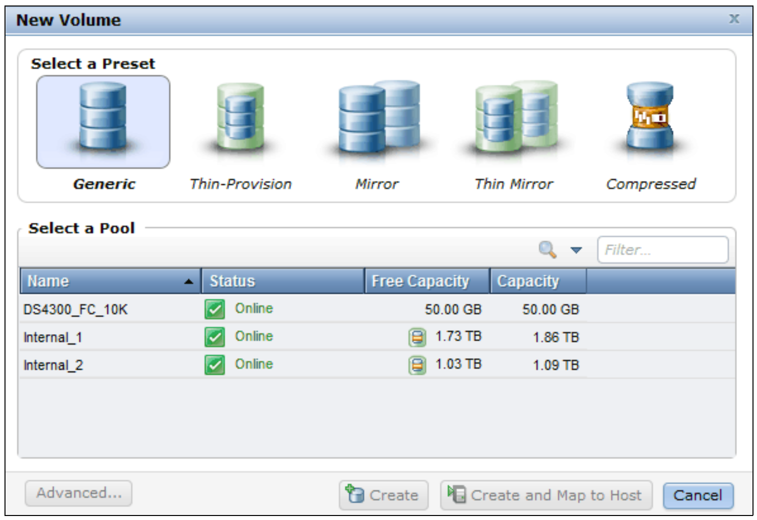


Figure 6-3 New Volume window - select pool

2. Select the managed disk pool where you want the volume to reside by clicking it. You may create multiple volumes by clicking the green plus sign below the Add Volumes column. We choose to create two new volumes. The result is shown in Figure 6-4.

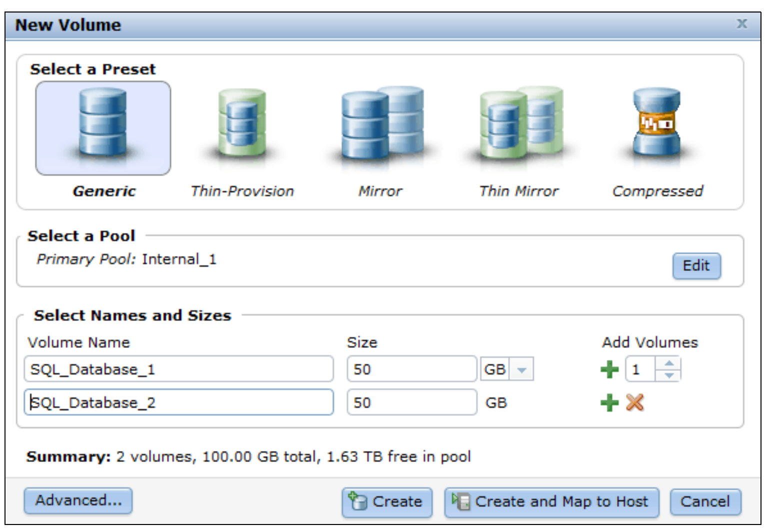


Figure 6-4 Create a generic volume

3. Enter a volume name and a size, and click **Create.** The new Generic volume is created and the CLI commands that IBM Flex System V7000 Storage Node uses are displayed (Figure 6-5 on page 245).

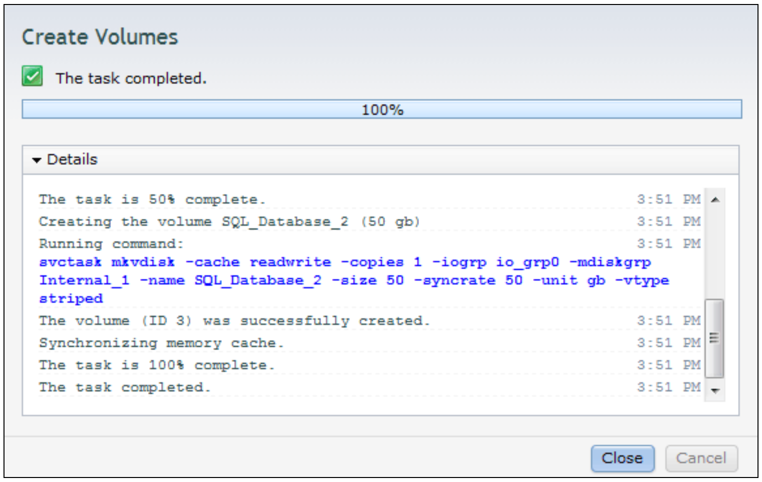


Figure 6-5 Volume created

4. After the task is completed, click **Close**

You can choose to map the new volumes immediately as part of the volume create wizard by clicking **Create and Map to Host**. This will guide you through the **Map to Host** wizard described in section 6.3, “Mapping a volume to the host” on page 261.

6.1.2 Creating a thin-provisioned volume

Volumes can be configured to be thin provisioned. A thin-provisioned volume behaves with respect to application reads and writes as though they were fully allocated. When creating a thin-provisioned volume, it is possible to specify two capacities: the real physical capacity allocated to the volume from the storage pool and its virtual capacity available to the host. So the real capacity determines the quantity of extents that are initially allocated to the volume. The virtual capacity is the capacity of the volume reported to all other components (for example, FlashCopy, cache, and Remote Copy) and to the host servers.

To create a thin-provisioned volume, complete the following steps:

1. From the New Volume window, select **Thin Provision** (Figure 6-6).

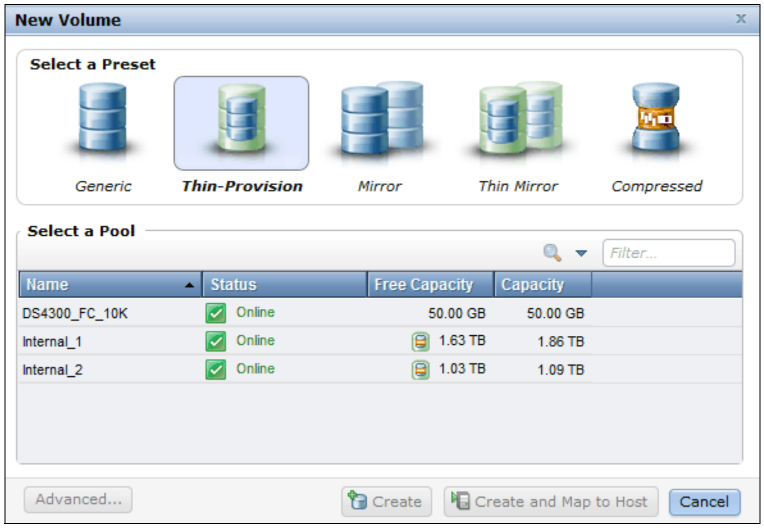


Figure 6-6 Create a thin-provisioned volume - select pool

2. Select the pool in which the thin-provisioned volume should be created by clicking it and entering the volume name and size (Figure 6-7).

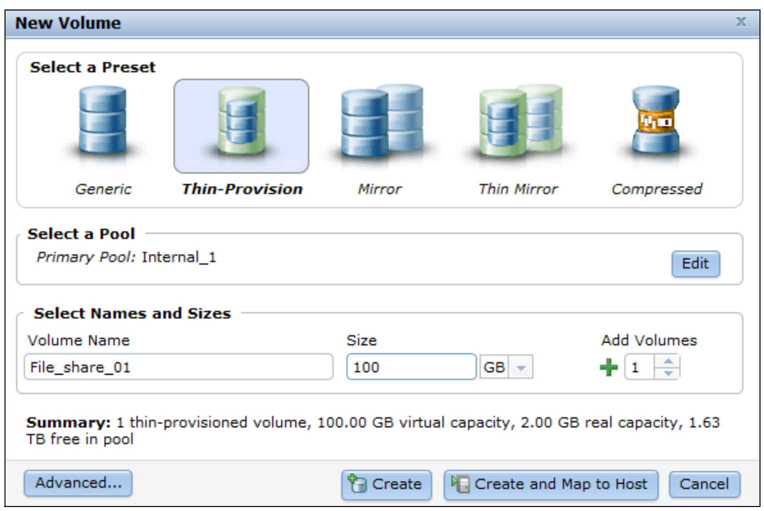


Figure 6-7 Enter the volume name and size

3. Under the Volume Name field is a summary showing that you are about to make a thin-provisioned volume, how much virtual space is will be available, the space that will be allocated (real size), and the free capacity in the pool. By default, the real capacity is 2% of the virtual capacity; you can change this setting by clicking **Advanced** and selecting **Capacity Management** as shown in Figure 6-8 on page 247.

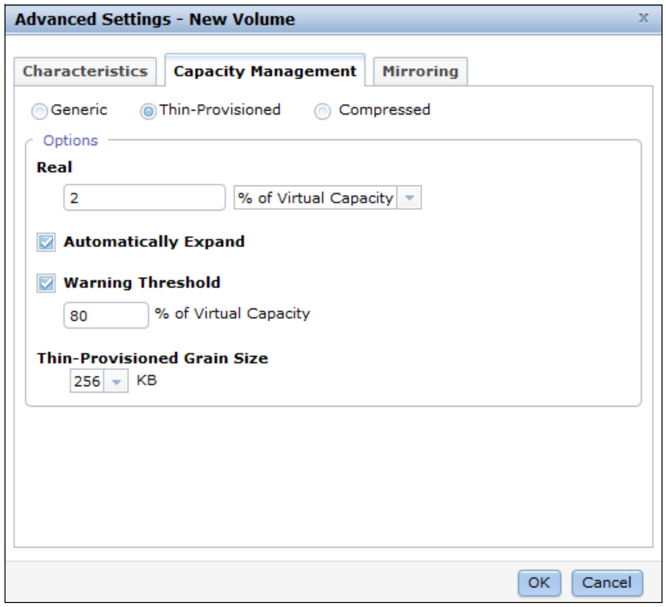


Figure 6-8 Advanced Settings

4. On the Thin Provisioning tab (Figure 6-8), there are several advanced options available:
   – Real: Specify the size of the real capacity space used during creation.
   – Automatically Extend: This option enables the automatic expansion of real capacity, if new capacity must be allocated.
   – Warning Threshold: Enter a threshold for receiving capacity alerts.
   – Thin Provisioned Grain Size: Specify the grain size for real capacity.

   Make your choices, if required, and click **OK** to return to Figure 6-7 on page 246.
5. In Figure 6-7 on page 246, click **Create** and the creation task completes as shown in Figure 6-9.

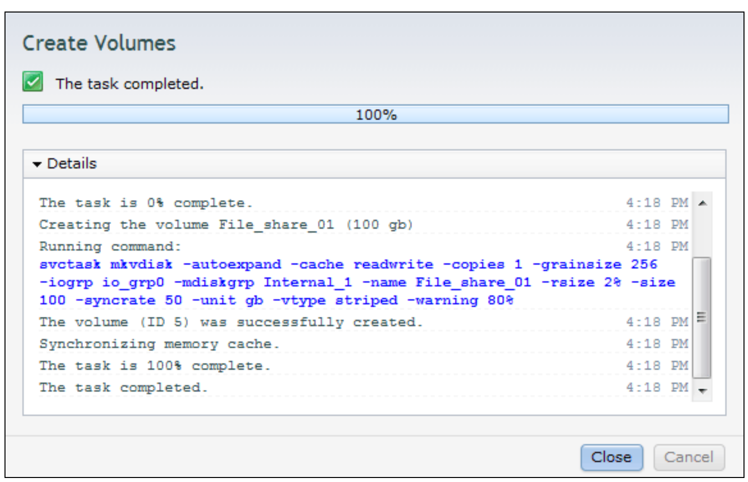


Figure 6-9 Create a thin-provisioned volume - CLI command displays

6. After the task is completed, click **Close**

You can choose to map the new volumes immediately as part of the volume create wizard by clicking **Create and Map to Host**. This will guide you through the **Map to Host** wizard described in section 6.3, "Mapping a volume to the host" on page 261.

6.1.3 Creating a mirrored volume

IBM Flex System V7000 Storage Node offers the capability to mirror volumes, which means a single volume is presented to the host, but two copies exist in the storage back end, usually in different storage pools (all reads are handled by the primary copy). This feature is similar to host based software mirroring such as LVM, but it provides a single point of management for all operating systems, and provides storage high availability to operating systems that do not support software mirroring.

This setup lets you protect against array failures (for example, multiple disk failures) and offers you more advanced features, as described in Chapter 9, "IBM Flex System V7000 Storage Node Copy Services" on page 341.It also allows you to greatly improve availability, but it is not a disaster recovery solution, due to both copies being accessed by the same node pair and only addressable by a single cluster.

In our example below we create a mirrored volume where the first copy, placed in the Primary Pool called *Internal_1*, is located internally to the IBM Flex System V7000 Storage Node disks. The second half of the mirror is placed in the Secondary Pool called *DS4300_FC_10K* and is placed externally to the IBM Flex System V7000 Storage Node on an IBM DS4300 disk array. By using this configuration the volume is tolerant to either internal or external disk array failure and hence does not have single point of failure.

For more information about storage pools and how to access external storage arrays refer to Chapter 8, "Storage pools" on page 295.

> **Note:** An unmirrored volume can be migrated from one location to another by simply adding a second copy to the desired destination, waiting for the two copies to synchronize, and then removing the original copy 0. This operation can be stopped at any time. The two copies can be in separate storage pools with separate extent sizes.

To create a mirrored volume, complete the following steps:

1. From the New Volume window, select **Mirror** (Figure 6-10).

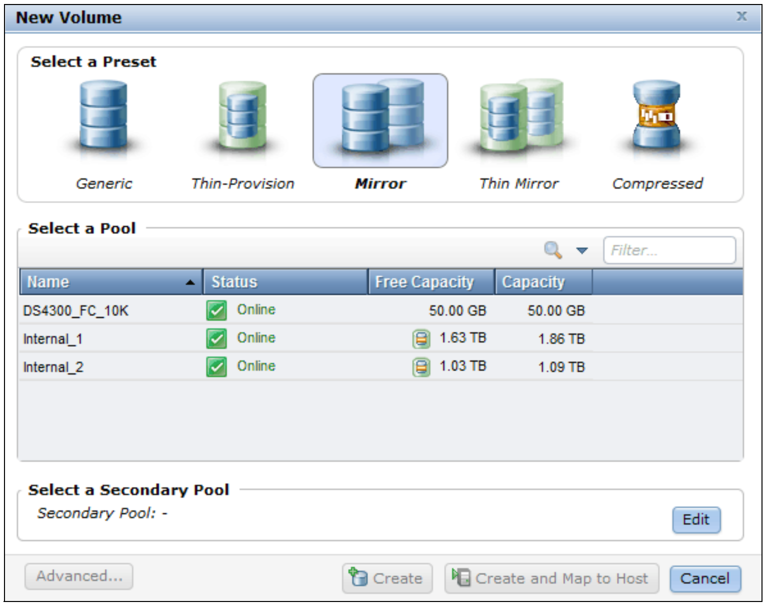


Figure 6-10 Create a mirrored volume - Select the primary pool

2. Select the primary pool by clicking it and the view changes to the second pool (Figure 6-11).

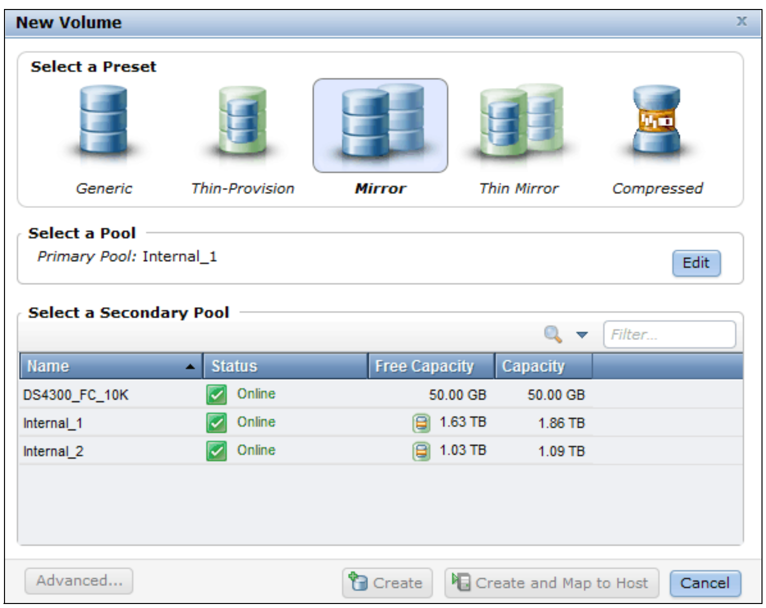


Figure 6-11 Create a mirrored volume - Select the secondary pool

3. Select the secondary pool by clicking it, and enter a volume name and the required size (Figure 6-12 on page 250).

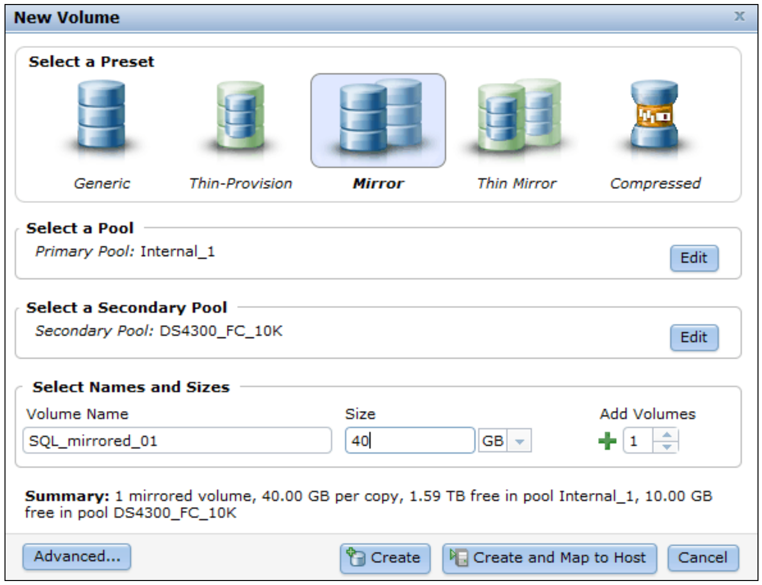


Figure 6-12 Create a mirrored volume - Select pools, name, and size

4. The summary shows you capacity information about the pool. If you want to select advanced settings, click **Advanced** and click the **Mirroring** tab (Figure 6-13).

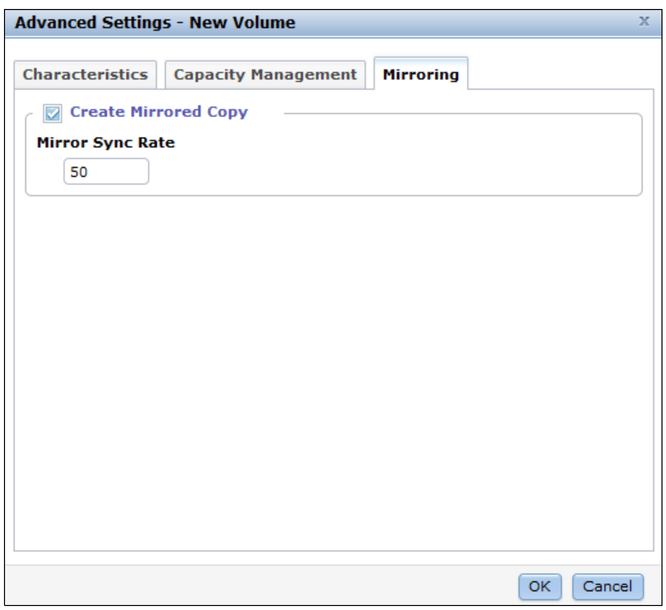


Figure 6-13 Advanced mirroring features

5. In the advanced mirroring settings, you are able to specify a synchronization rate. Enter a Mirror Sync Rate between 1 - 100%. With this option, you can set the importance of the copy synchronization progress, which enables you to prefer more important volumes for synchronization faster than other mirrored volumes. By default, the rate is set to 50% for all volumes. Click **OK** to return to Figure 6-12.

6. Click **Create and Map to Host** and the mirrored volume is created (Figure 6-14).

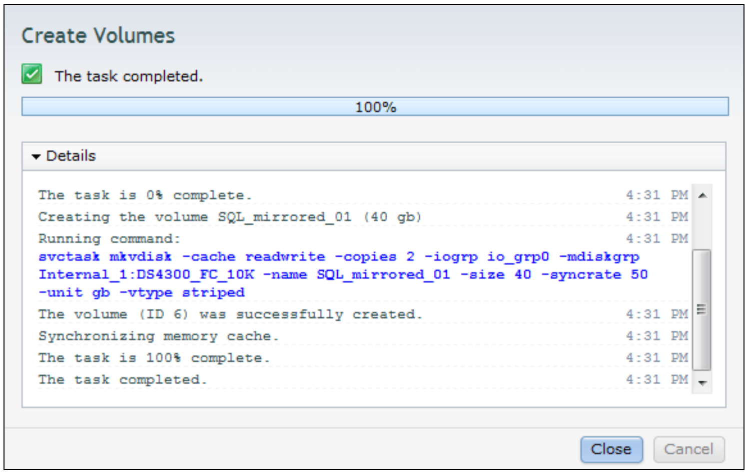


Figure 6-14 Create a mirrored volume - CLI command displays

7. After the task is completed, click **Close**

You can choose to map the new volumes immediately as part of the volume create wizard by clicking **Create and Map to Host**. This will guide you through the **Map to Host** wizard described in section 6.3, “Mapping a volume to the host” on page 261.

6.1.4 Creating a thin-mirror volume

A thin-mirror volume allows you to allocate the required space on demand (described in 6.1.2, “Creating a thin-provisioned volume” on page 245) and have several copies of a volume (described in 6.1.3, “Creating a mirrored volume” on page 248) available.

In the example below we create a thin-mirror volume where one copy of the mirror is physically located internally to the IBM Flex System V7000 Storage Node disks and the other copy is located on an IBM DS4300 disk array system.

For more information about storage pools and how to access external storage arrays refer to Chapter 8, “Storage pools” on page 295.

To create a thin-mirror volume, complete the following steps:

1. From the New Volume window, select **Thin Mirror** (Figure 6-15).

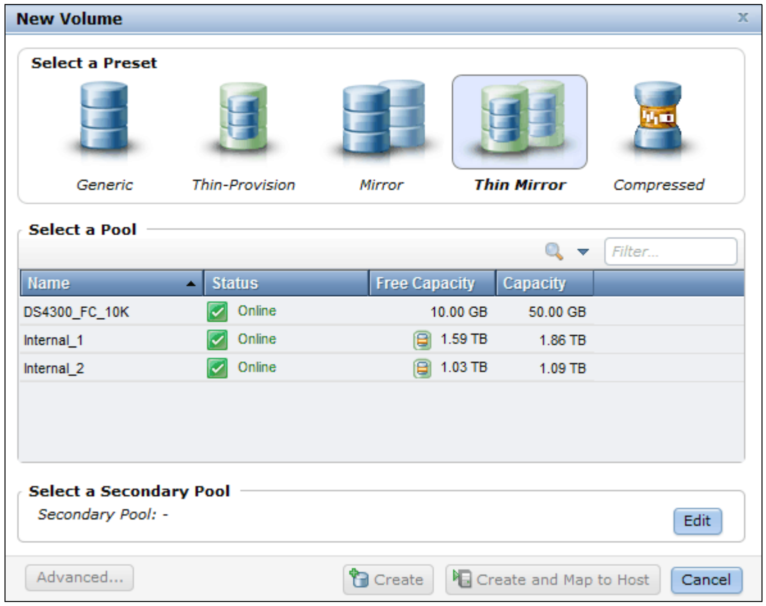


Figure 6-15 Thin Mirror - select the primary pool

2. Select the primary pool by clicking it and the view changes to the second pool (Figure 6-16).

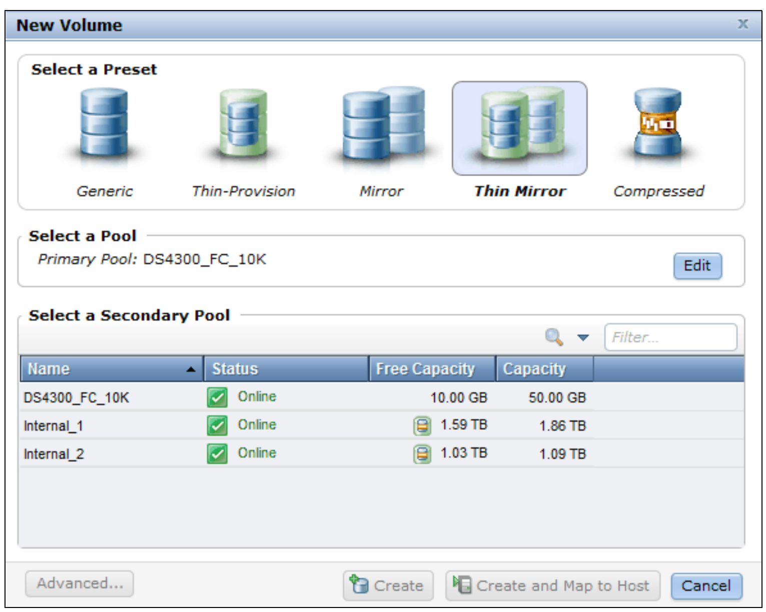


Figure 6-16 Thin Mirror - select the secondary pool

3. Select the pool for the secondary copy and enter a name and a size for the new volume (Figure 6-17 on page 253).

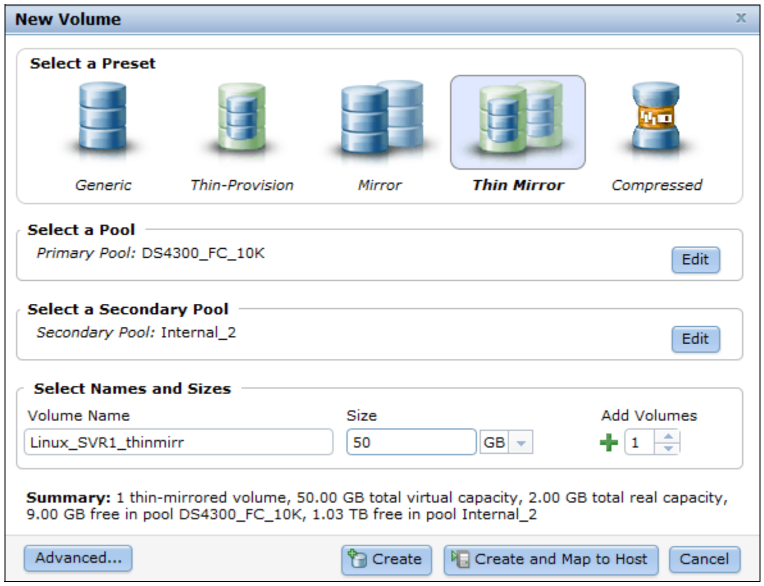


Figure 6-17 Thin Mirror - Enter volume name and size

4. The summary shows you the capacity information and the allocated space. You can click **Advanced** and customize the thin-provision settings (Figure 6-8 on page 247) or the mirror synchronization rate (Figure 6-13 on page 250). If you have opened the advanced settings, click **OK** to return to Figure 6-17.
5. Click **Create** and the mirrored volume is created (Figure 6-18).

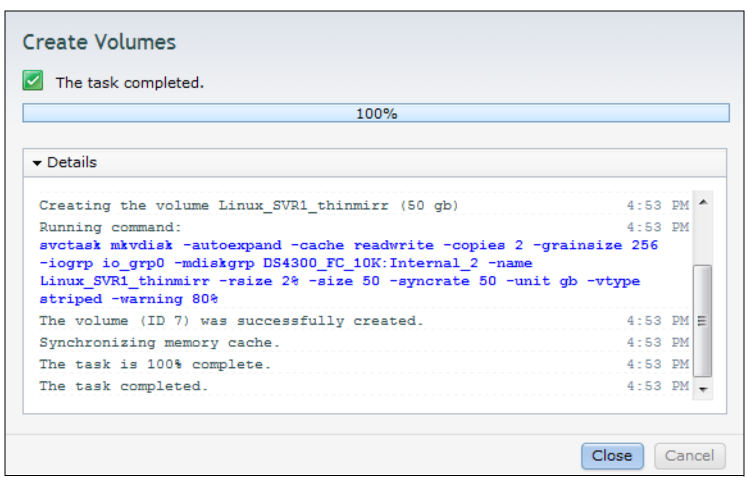


Figure 6-18 Thin Mirror volume creates - CLI command displays

6. After the task is completed, click **Close**

You can choose to map the new volumes immediately as part of the volume create wizard by clicking **Create and Map to Host**. This will guide you through the **Map to Host** wizard described in section 6.3, “Mapping a volume to the host” on page 261.

6.1.5 IBM Real-time Compression

IBM Flex System V7000 version 6.4 (or higher) has the capability of compressing volumes and this is called IBM Real-time Compression. This reduces disk capacity required to store data by up to 80%, helping improve efficiency and control storage costs.

The IBM Real-time Compression solution addresses the requirement for reducing IT costs by reducing the amount of physical capacity required and by reducing the required floorspace, power and cooling. Implementing Real-time Compression in IBM Flex System V7000 Storage Node provides the following benefits:

- Compression for active primary data:

  IBM Real-time Compression can be used with active primary data, therefore it supports workloads which are not candidates for compression in other solutions. The solution supports online compression of existing data and allows storage administrators to gain back free disk space in the existing storage system without enforcing administrators and users to clean up or archive data. This significantly enhances the value of existing storage assets and the benefits to the business are immediate. Capital expense of upgrading or expanding the storage system is delayed.

- Compression for replicated/mirrored data:

  Remote volume copies can be compressed in addition to the volumes at the primary storage tier. This reduces storage requirements in Metro Mirror and Global Mirror destination volumes as well.

- No changes to the existing environment are required:

  IBM Real-time Compression is an integral part of the storage system, and was designed with transparency in mind so that it can be implemented without changes to applications, Creating a Compressed Volume

Creating a compressed volume

To create a Compressed volume, complete the following steps:

1. Click **New Volume** and select **Compressed** as shown in Figure 6-19.

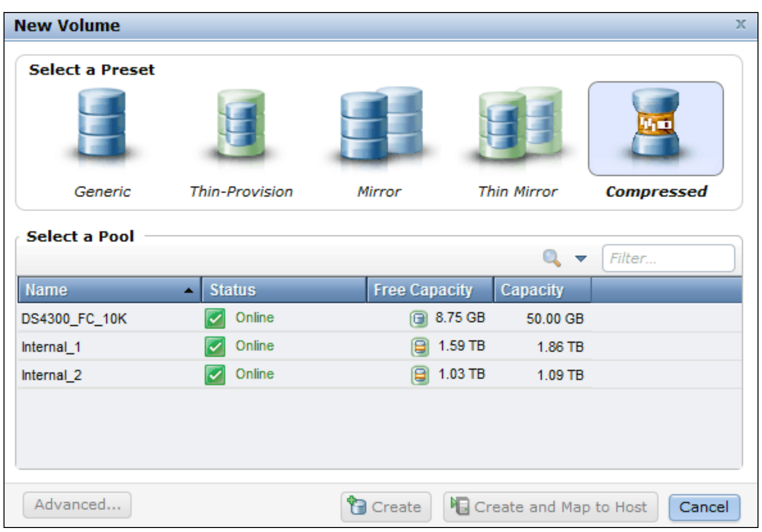


Figure 6-19 Compressed Volume - select a pool

Note that a warning is given that creating a compressed volume may affect performance of existing volumes. Hence creating compressed volumes may require some planning in case

this is a very busy system. One option is to create the compressed volume at the end of the day, so that fewer users are affected by a performance decrease during creation of the compressed volume.

2. Select a pool where you want the new Compressed volume as shown in Figure 6-20. Type a name for the new Compressed volume, and the size and click Create.

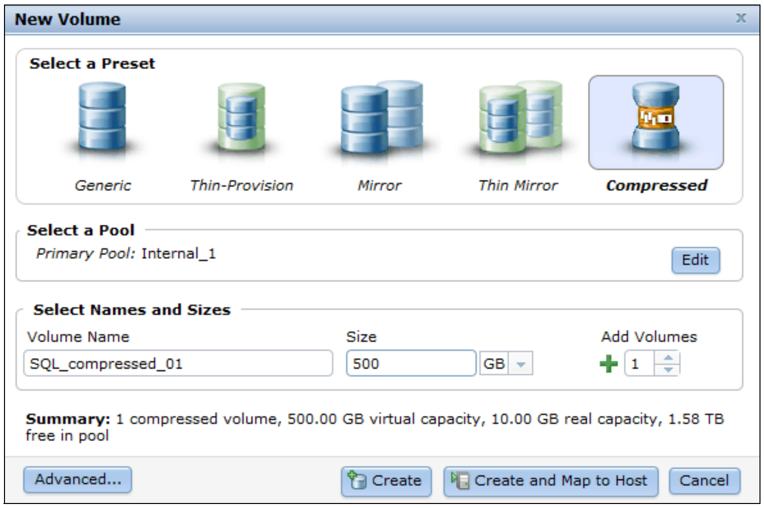


Figure 6-20 Compressed Volume - Enter name and size

3. In Figure 6-20, you can click Advanced and choose capacity management options as shown in Figure 6-21.

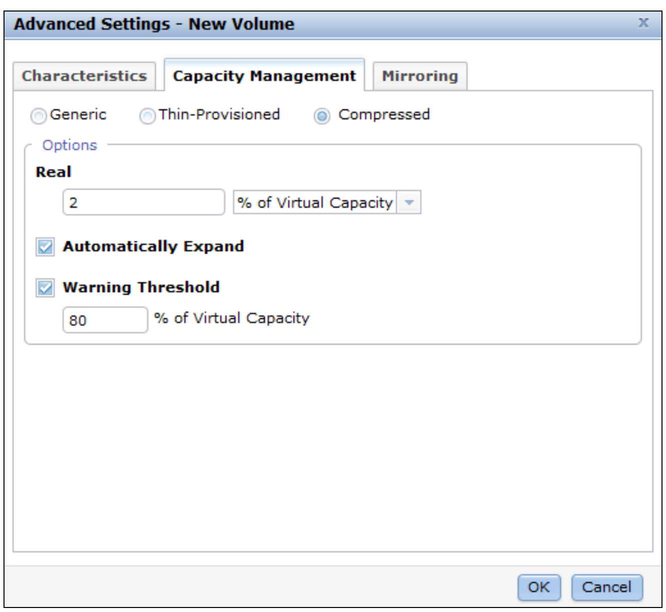


Figure 6-21 Compressed Volume - Advanced properties

4. The Compressed volume creates while the CLI command displays Figure 6-22 on page 256.

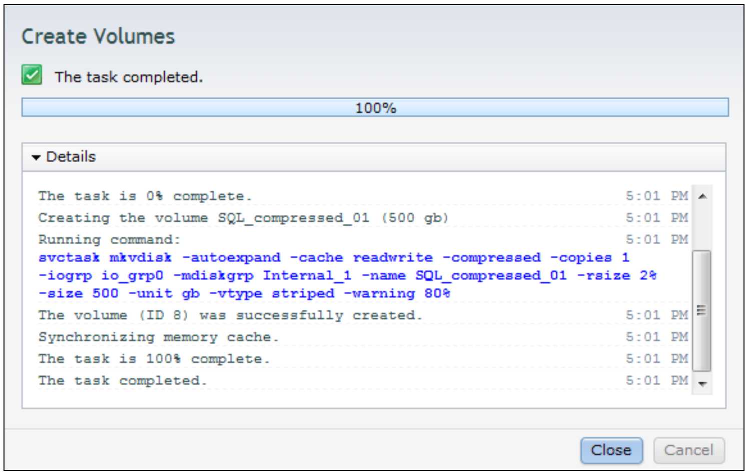


Figure 6-22 Compressed Volume- CLI command displays

5. After the task is completed, click **Close**

You can choose to map the new volumes immediately as part of the volume create wizard by clicking **Create and Map to Host**. This will guide you through the **Map to Host** wizard described in section 6.3, “Mapping a volume to the host” on page 261.

For more information about Real-time Compression in IBM Flex System V7000 Storage Node as well as IBM Storwize V7000 and IBM SAN Volume Controller check *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859.

6.2 Creating a new host

Before we can map any volumes to a host, the host has to be created first. We here demonstrate how to create new hosts in the IBM Flex System V7000 Storage Node.

6.2.1 Create a Fibre Channel attached host

To create the Fibre Channel host do the following:

1. Navigate to the **Hosts** tab of the IBM Flex System V7000 Storage Node GUI and click **New Host** as shown in Figure 6-23 on page 257.

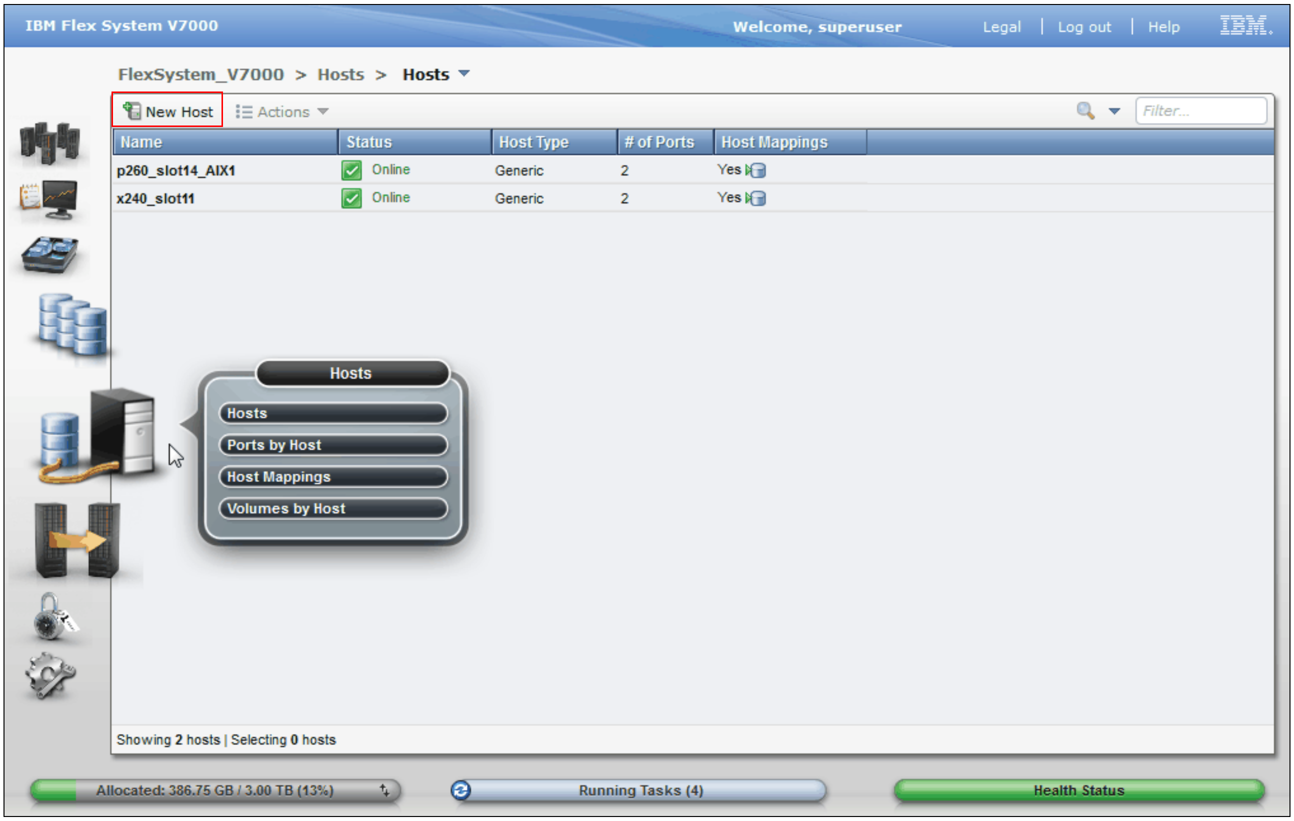


Figure 6-23 The Hosts menu - create new host

2. Click **Fibre Channel Host** - the Wizard automatically moves to the next screen.

   Figure 6-24 shows that the first choice we have to make is whether we are creating a Fibre Channel attached host or an iSCSI host.

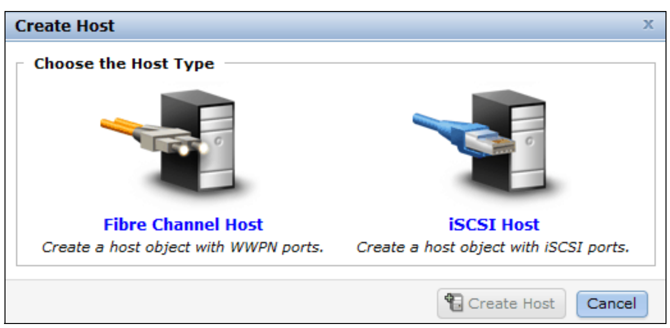


Figure 6-24 New Host creation - select iSCSI or FC host

3. Give the host a name after your choice of the host type.
4. Select the World Wide Names (WWNs) that represent the Host Bus Adaptor (HBA) ports in your server. If these do not appear on the list click **Rescan**. If they still not appear check your SAN-switch Zoning.

   Figure 6-25 on page 258 shows that we are selecting the two WWNs that represent the two HBAs in our server.

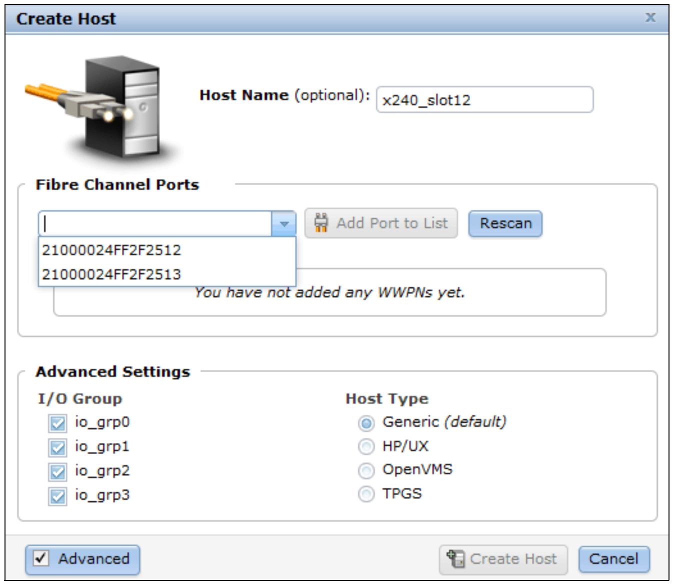


Figure 6-25 New Host creation - select between available World Wide Names

5. Click **Create Host** after you finished adding WWNs

   Figure 6-26 shows that we have added our HBA ports and are now ready to create our new host.

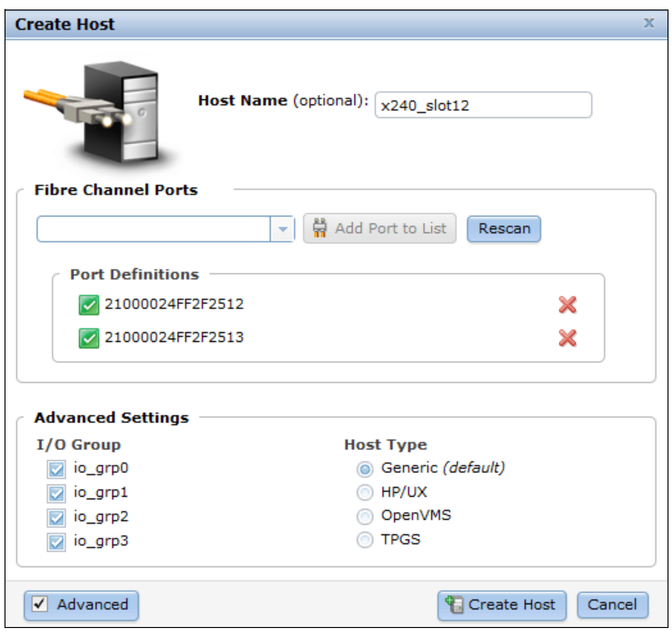


Figure 6-26 New Host creation - two World Wide Names Selected

The CLI commands window appear and after the task is completed, click **Close.** The FC-attached host is now created successfully and it now appears on the list of configured hosts. You can right click on the server here to perform any additional configuration tasks as shown in Figure 6-27 on page 259.

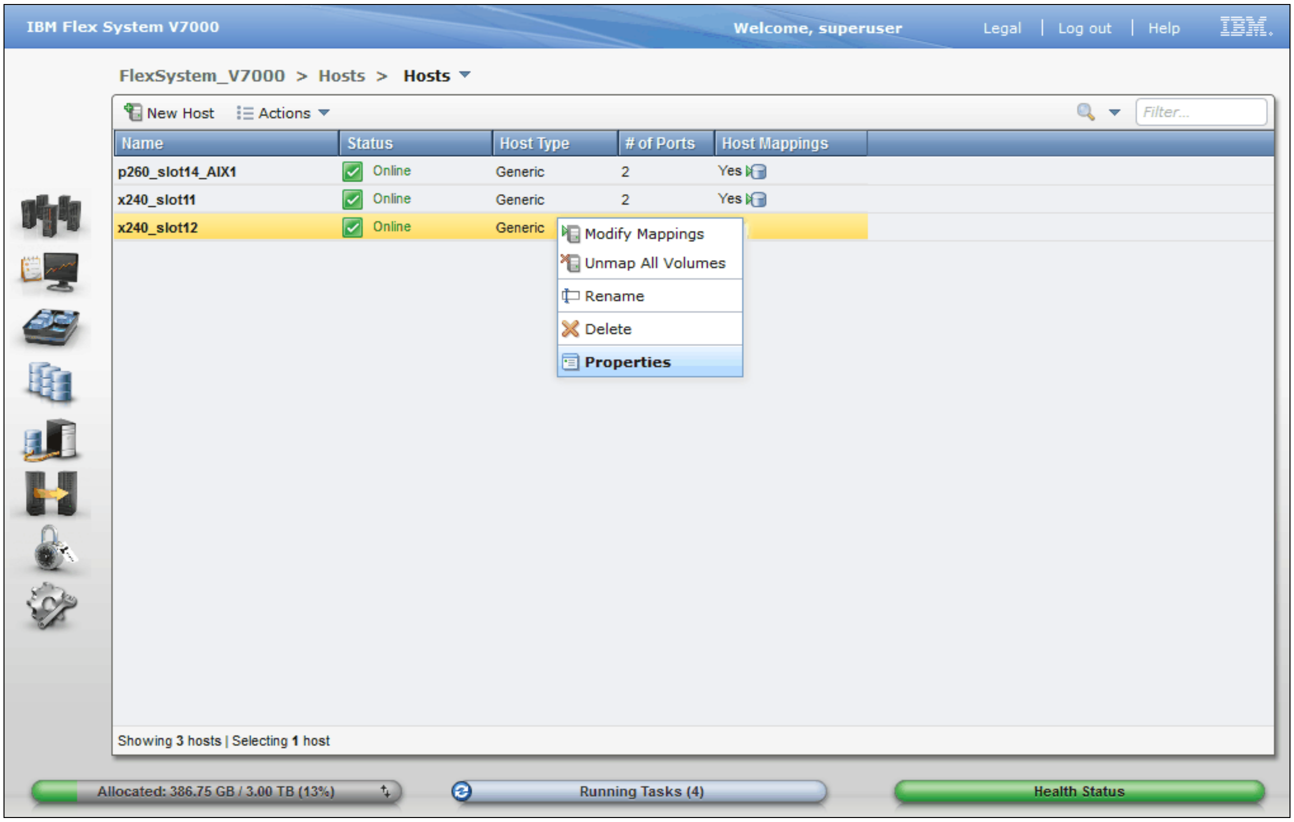


Figure 6-27 The newly created host appears

6.2.2 Create an iSCSI attached host

The process of creating a iSCSI attached host is almost similar to creating a FC-attached host as shown in 6.2.1, "Create a Fibre Channel attached host" on page 256. The only difference is that instead of configuring the new host with WWN's from our host we instead configure an iSCSI initiator name. The iSCSI initiator name is obtained from the iSCSI initiator software installed on the host computer and represents a unique - however changeable iSCSI address of our host on the network.

To create the iSCSI host do the following:

1. As described in 6.2.1, "Create a Fibre Channel attached host" on page 256 navigate to **Hosts** and click **New Host**.
2. Select iSCSI host
3. At the Create Host windows type a name for the new host
4. Insert the iSCSI initiator name which you copied from the hosts iSCSI initiator software.

Figure 6-28 on page 260 shows that we pasted in the iSCSI name of the new host.

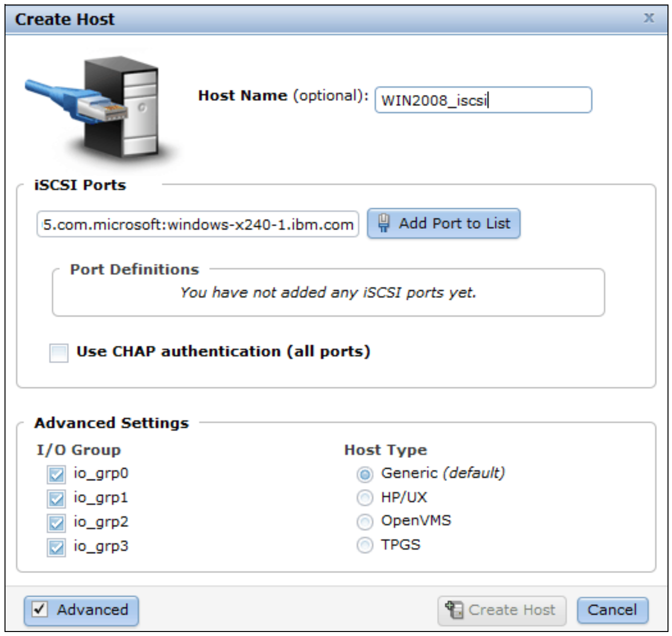


Figure 6-28 New Host creation - add iSCSI port name obtained from host

5. Click **Create Host.** Figure 6-29 shows the final step of creating the iSCSI host.

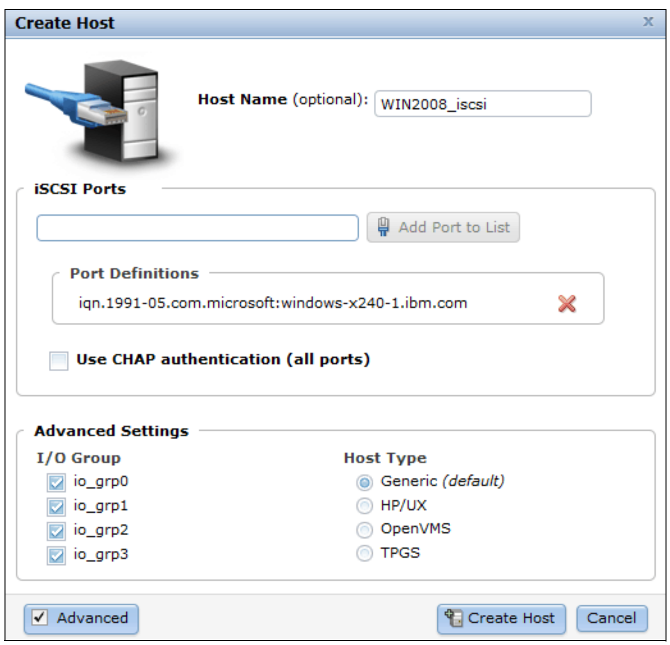


Figure 6-29 New Host creation - iSCSI port defined

Figure 6-30 on page 261 shows that we now have two configured hosts in our IBM Flex System V7000 Storage Node GUI.

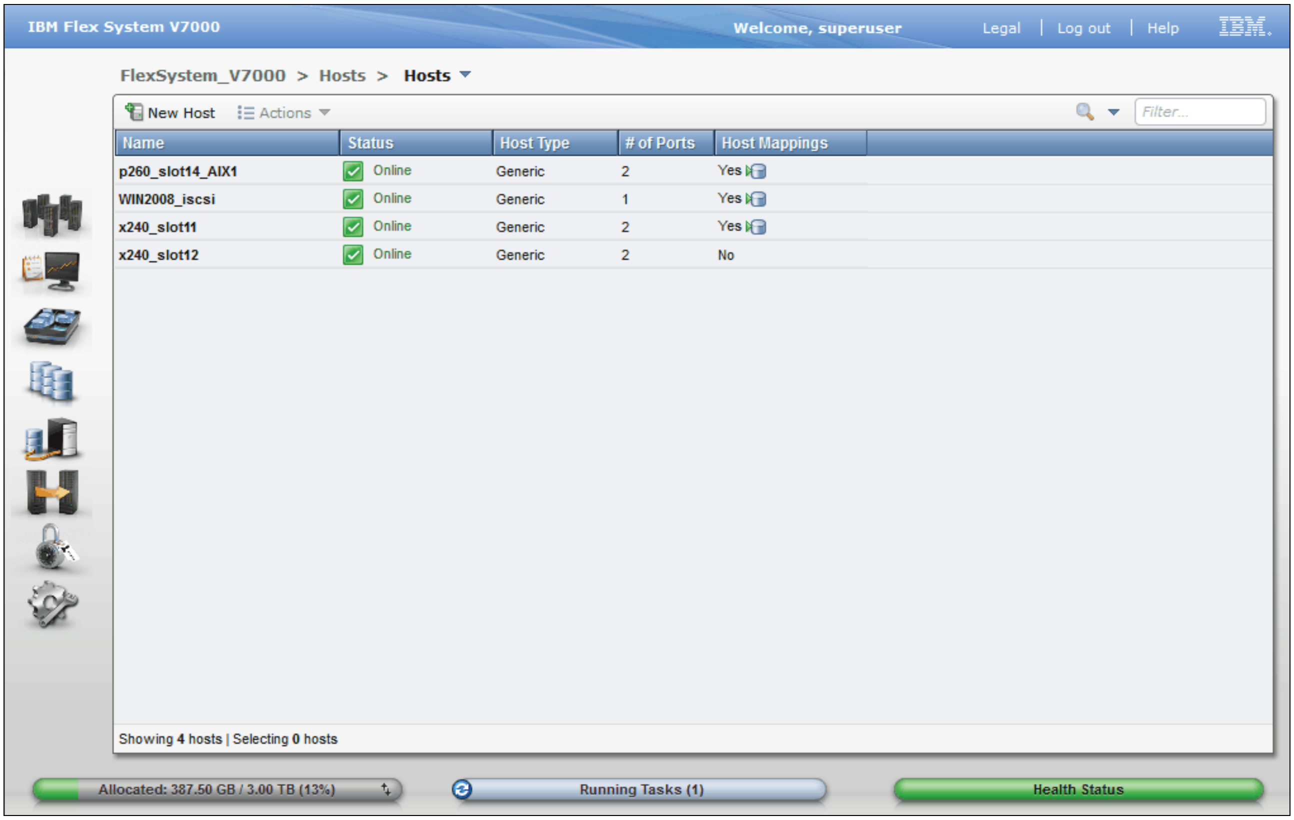


Figure 6-30 The Hosts menu - two additional hosts created

We have now finished creating the Fibre Channel attached host as well as the iSCSI attached host. The newly created iSCSI host may show **Status** with the color red and **Offline**. This indicates that no traffic has yet travelled over the connection. To make it green and online you have to connect to the iSCSI attached host and from there discover the IBM Flex System V7000 Storage Node portal.

A single host can contain both FC-ports as well as iSCSI ports, but in our example we created two different hosts in the host section of IBM Flex System V7000 Storage Node GUI representing the same physical server.

Note: The iSCSI host will not show online ports before the connection has been discovered from the host which makes the connection active.

6.3 Mapping a volume to the host

In this section we describe how to map a volume to a host. The map to host wizard either comes after the volume create wizard or happens during.

6.3.1 Mapping newly created volumes to the host using the wizard

In section 6.1, “Storage provisioning from IBM Flex System V7000 Storage Node” on page 242, we created the various type of volumes available on IBM Flex System V7000 Storage Node. The volume creation wizard gives the option to **Create** or **Create and Map to Host**.

If you clicked **Create** the next logical step is to map the newly created volume to a host. If you clicked **Create and Map to Host** the volume create wizard takes you to the host mapping wizard as shown below.

To map the volumes, complete the following steps:

1. From the volumes menu select a volume and click either **Actions** and **Map to Host** or right-click and select **Map to host** as shown in Figure 6-31.

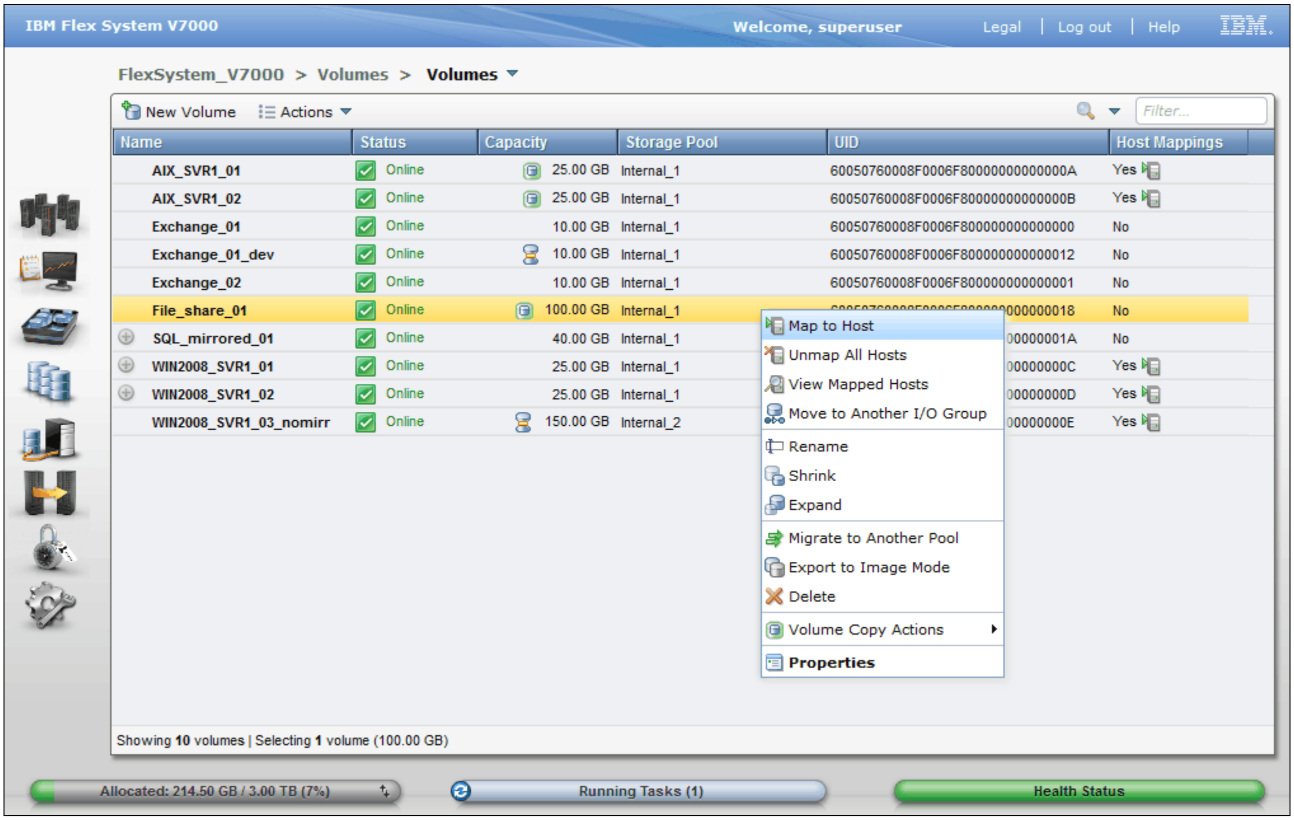


Figure 6-31 The volumes menu - map to host

2. Select a host to which the new volume should be attached (Figure 6-32).

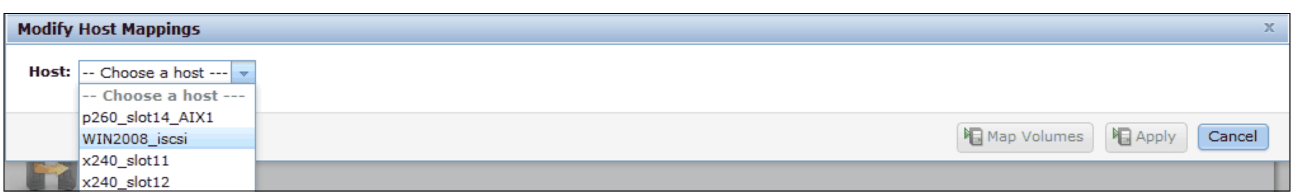


Figure 6-32 Map to host - choose host

3. The Modify Mappings window opens, and your host and the newly created volume are already selected. Click **OK** and the volume is mapped to the host (Figure 6-33 on page 263).

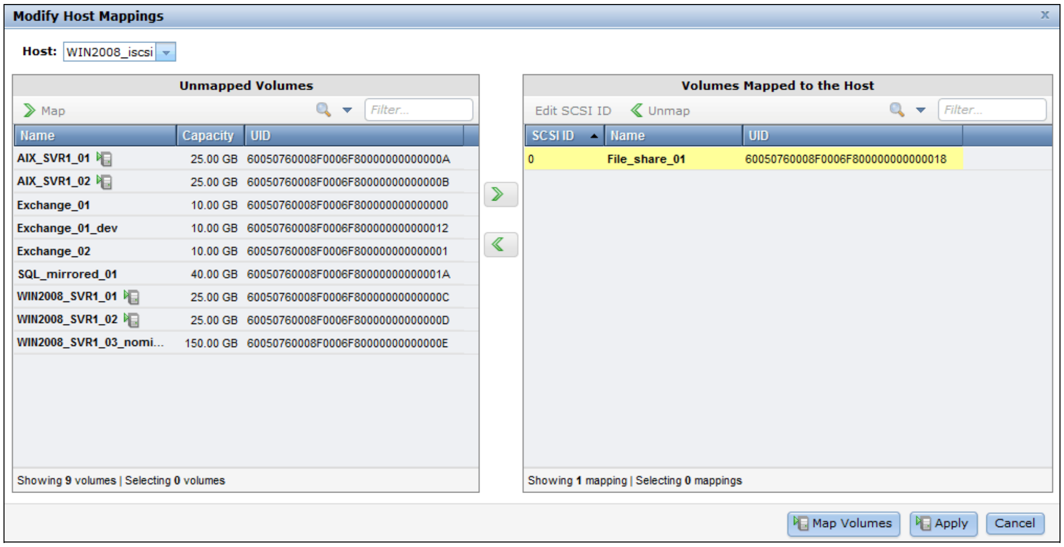


Figure 6-33 Map to host - select volumes

4. After the task completes, click **Close** (Figure 6-34), and the wizard returns to the All volumes window.

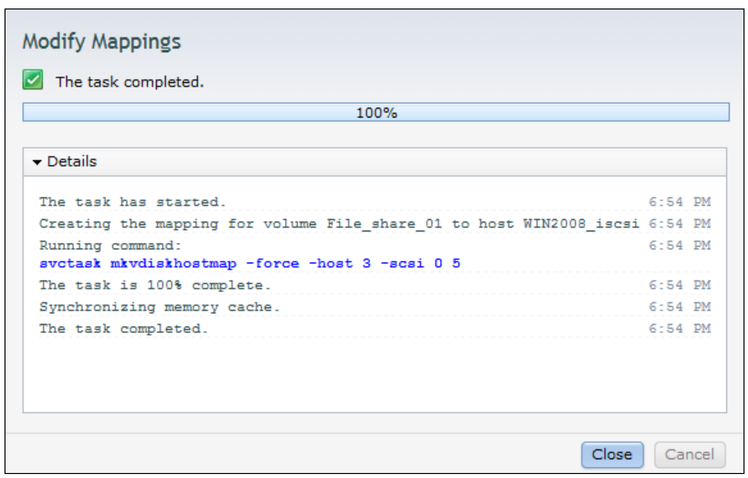


Figure 6-34 Map to host - CLI command displays

The newly created volume is displayed and we see that it is now has host mappings as shown in Figure 6-35 on page 264.

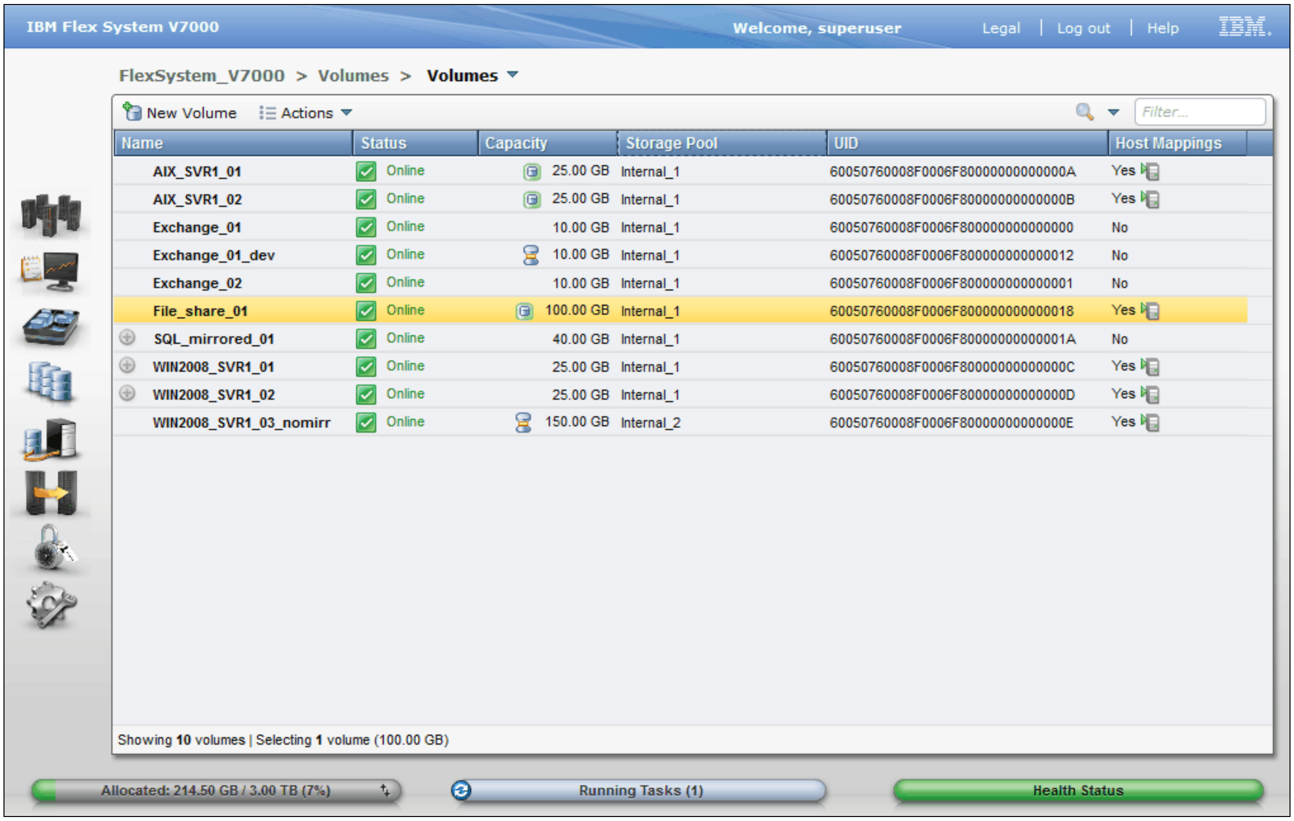


Figure 6-35 Map to host - Volume now has host mappings

The host is now able to access the volumes and store data on it. Go to Chapter 12, “Host Configuration” on page 473 to discover the volumes on the host and make some additional host settings if required.

You can also create multiple volumes in preparation for discovering them later. Mappings can be customized as well. For information about how to prepare hosts for connecting to the IBM Flex System V7000 Storage Node refer to Chapter 12, “Host Configuration” on page 473.

7

Migration wizard

In this chapter, we describe how to migrate existing data on external storage systems to the IBM Flex System V7000 Storage Node. Migrating data from external storage systems to the IBM Flex System V7000 Storage Node allows applications to benefit from the new features, such as Easy Tier, an intuitive management GUI, and advanced storage replication functions that better support applications.

To migrate existing data, the IBM Flex System V7000 Storage Node provides a storage migration wizard to guide you through the entire procedure.

7.1 Preparing for data migration

In this section, we show you how to prepare external storage systems and IBM Flex System V7000 Storage Node for data migration. Data migration is always an important step when setting up new storage systems in an existing storage infrastructure.

Before attaching any external storage systems to IBM Flex System V7000 Storage Node, see the IBM Flex System V7000 Storage Node support matrix at the following address:

http://www-947.ibm.com/support/entry/portal/overview/hardware/puresystems/pureflex_system/storage_node/flex_system_v7000

Also always check for supported hardware and driver versions at IBM System Storage Interoperation Center (SSIC):

http://ibm.com/systems/support/storage/ssic/interoperability.wss

Also available is the *IBM Flex System Interoperability Guide*, REDP-FSIG which is a Quick reference for IBM Flex System Interoperability.

When migrating data from an external storage system to the IBM Flex System V7000 Storage Node, where the external storage system is removed when complete, you can temporarily configure the external virtualization license setting. Configuring the external license setting prevents messages from being sent that indicate that you are in violation of the license agreement. When the migration is complete, the external virtualization license must be reset to its original limit.

> **External storage:** When external storage is virtualized by IBM Flex System V7000 Storage Node, a per-enclosure external virtualization license is required. Contact your IBM account team or IBM Business Partner for further assistance, if required.

To prepare the data migration, external storage systems need to be configured to be under IBM Flex System V7000 Storage Node control. Complete the following steps:

1. Stop host I/O to the external storage LUNs that need to be migrated.
2. Remove zones between the hosts and the storage system from which you are migrating.
3. Update your host device drivers, including your multipath driver, and configure them for attachment to the IBM Flex System V7000 Storage Node system.
4. Create a storage system zone between the storage system being migrated and IBM Flex System V7000 Storage Node system, and the host zones for the host attachment.
5. Unmap the LUNs in the external storage system to the host and map them to the IBM Flex System V7000 Storage Node system.
6. Verify that the IBM Flex System V7000 Storage Node has discovered the LUNs as unmanaged MDisks.

7.2 Migrating the data using the migration wizard

In this section, we describe how to migrate existing data on external storage systems using the IBM Flex System V7000 Storage Node storage migration wizard.

We use the IBM Flex System V7000 Storage Node wizard that has been designed specifically for this scenario to guide you through the process and describe the steps along the way.

7.2.1 Checking the Windows 2008 host before upgrade

Before any migration begins at the IBM Flex System V7000 Storage Node Migration Wizard we check the host that currently has its disks located on a IBM DS4300 disk array.

Figure 7-1 shows that the server has a disk two disks (Disk 1 and Disk 2) which are currently located on a IBM DS4300 disk array.

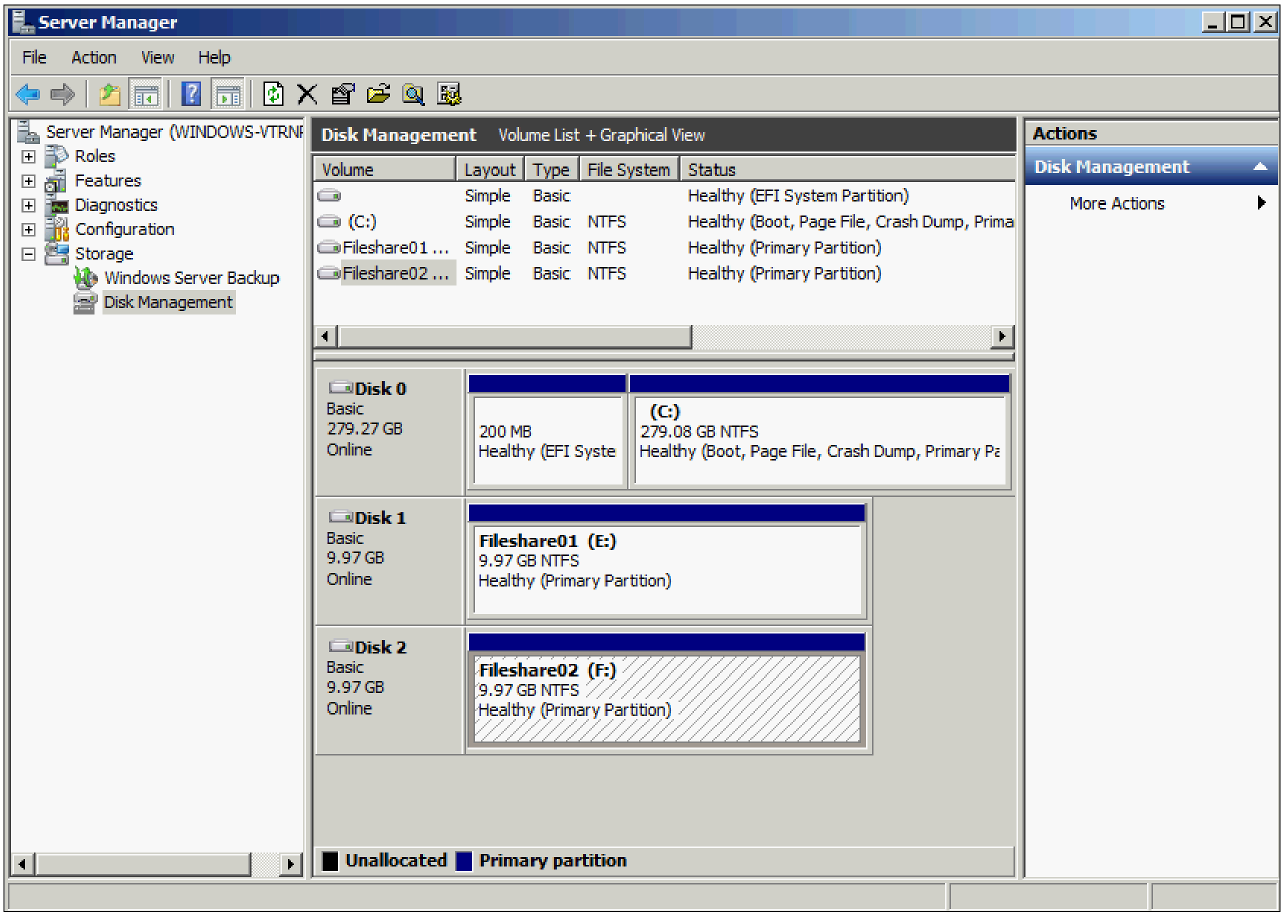


Figure 7-1 Windows server with disks from DS4300 mapped

7.2.2 Remapping the disk to the IBM Flex System V7000 Storage Node

Next we jump to the IBM System Storage DS® Storage Manager from where we unmap the disk on the DS4300 and remap it to the IBM Flex System V7000 Storage Node.

At this point we have two options:

- In the Server Manager Disk Administrator, take the disk offline, so that the Microsoft Windows 2008 server does not loose connectivity to an online disk
- Shutdown the Microsoft Windows 2008 server. The server may also need to be rebooted after installing the SDDDSM drivers which provides Multipathing support for IBM Flex System V7000 Storage Node.

We choose to shut down the server while remapping the DS4300 LUNs.

Figure 7-2 on page 268 shows that the DS Storage Manager is currently mapping the disk to the Windows server.

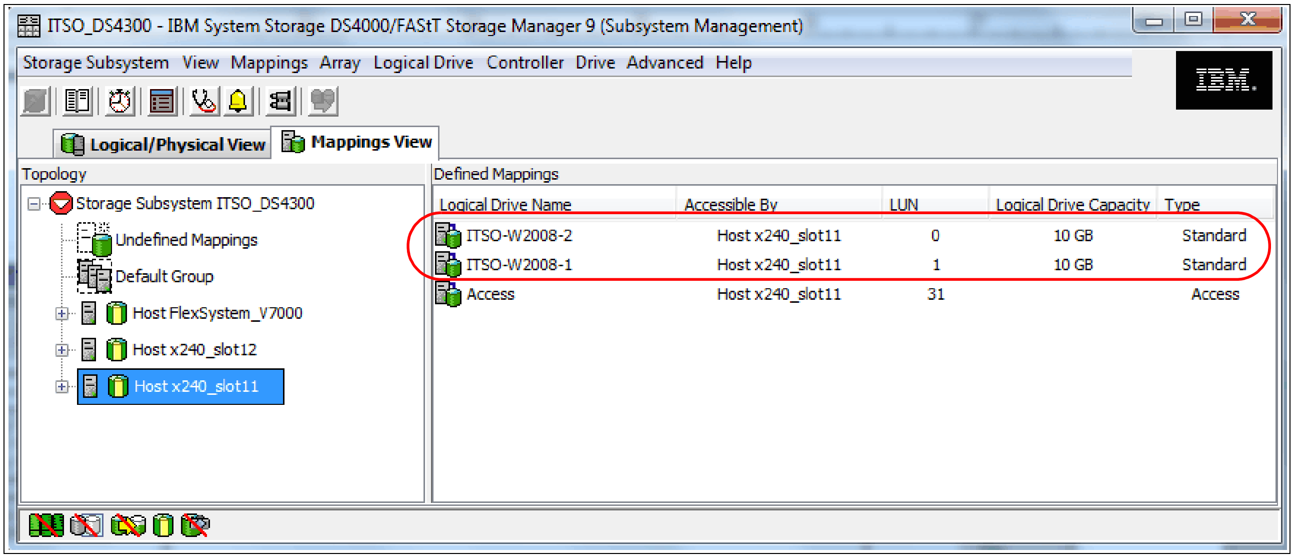


Figure 7-2 DS4300 Storage Manager - Disk maps to Windows server

Figure 7-3 shows how we make the mapping changes in DS Storage Manager. The disk is unmapped from the Windows server and is mapped to the IBM Flex System V7000 Storage Node instead.

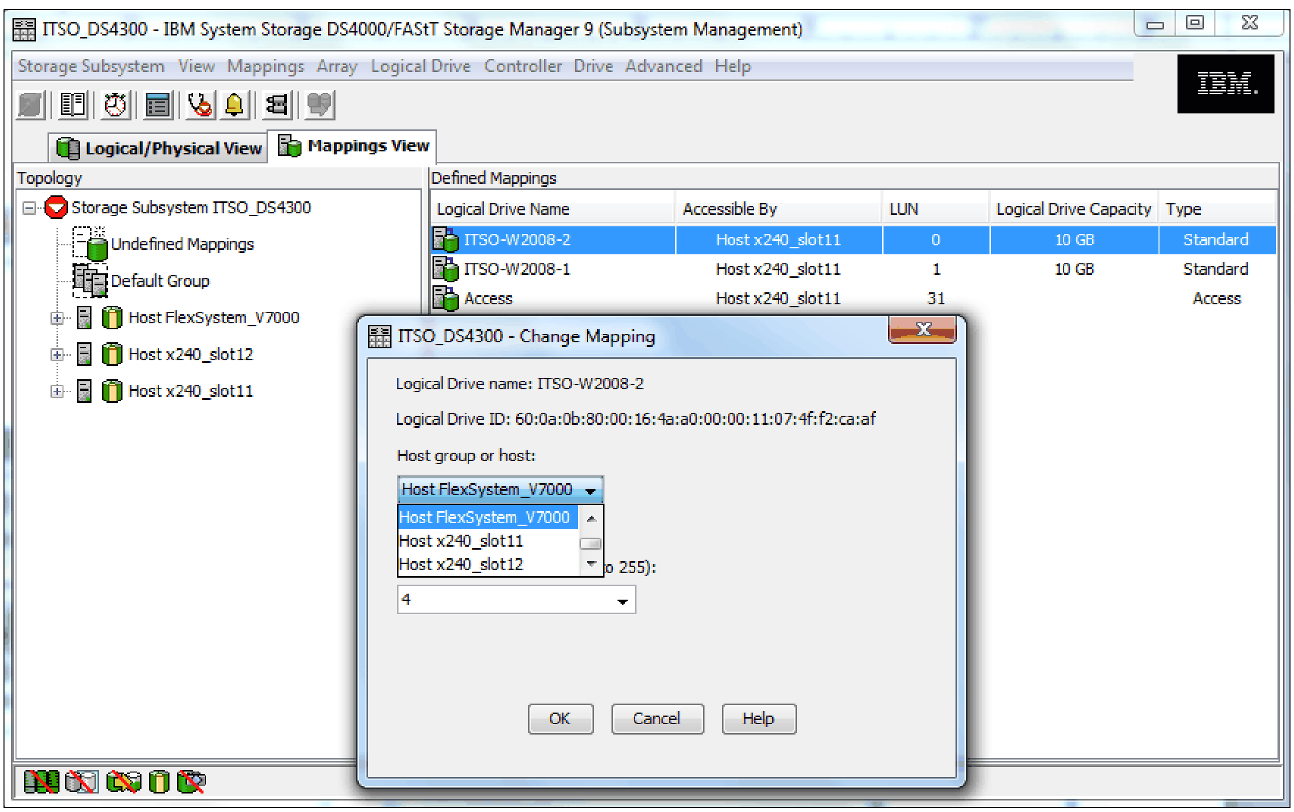


Figure 7-3 DS4300 Storage Manager - Disk Remapping

Figure 7-4 on page 269 shows that the disks are no longer mapped to the Windows server but are mapped to the IBM Flex System V7000 Storage Node.

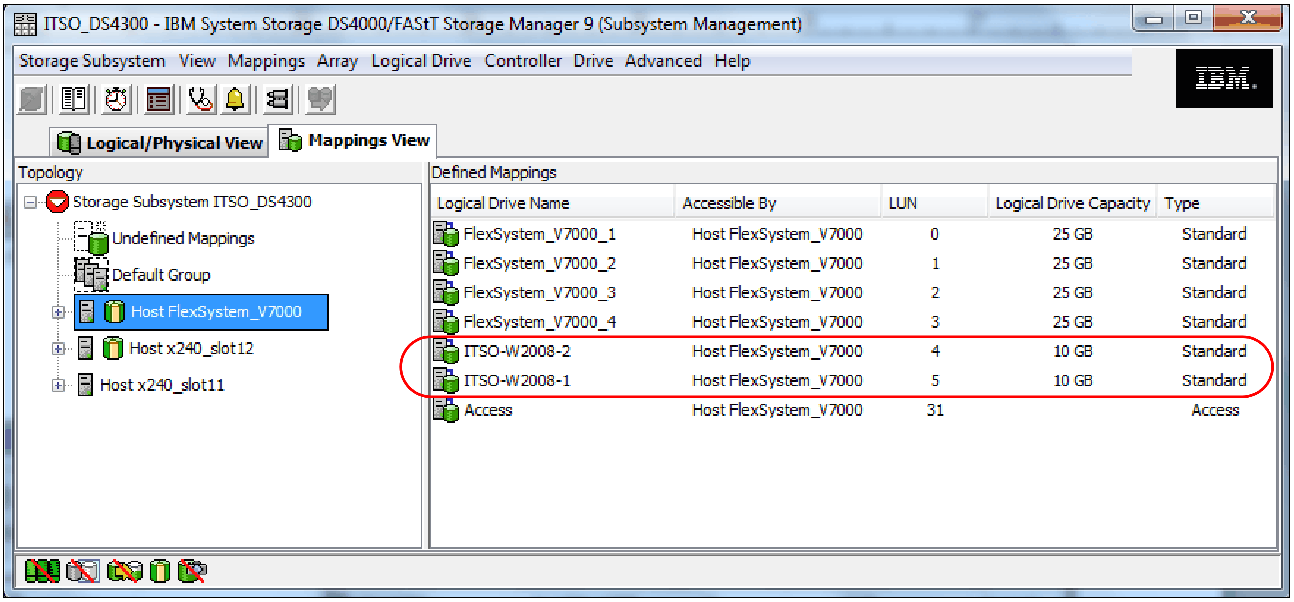


Figure 7-4 DS4300 Storage Manager - Disk now maps to IBM Flex System V7000 Storage Node

At this point we go back to our Microsoft Windows 2008 server and check that the disk has disappeared from the Server Manager Disk Administrator. Figure 7-5 on page 270 shows that the disk has disappeared from the Server Manager Disk Administrator. We also use the Rescan Disks option to confirm this.

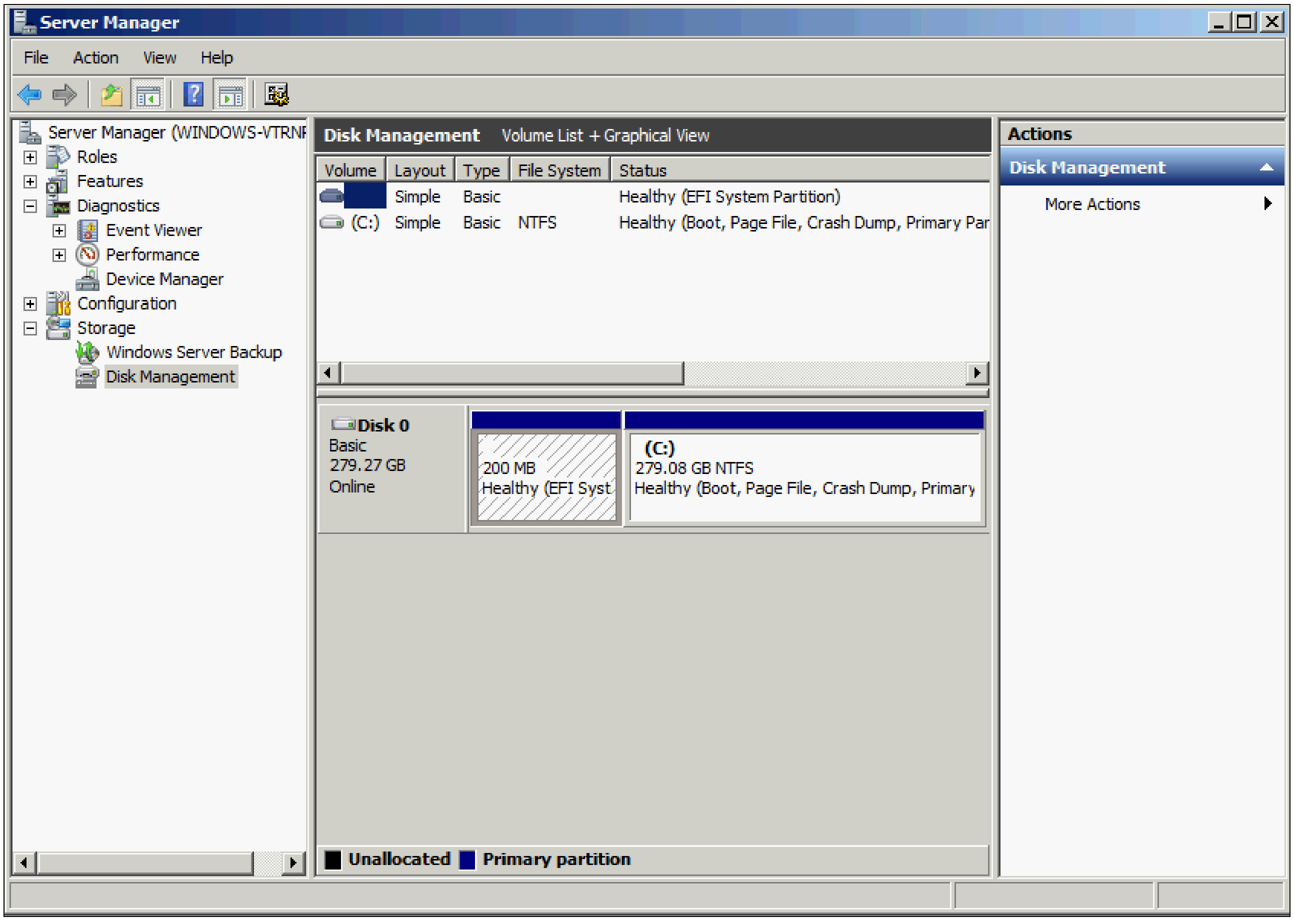


Figure 7-5 Windows server with disk from DS4300 unmapped - rescan disks

Now we are ready to start the Import of the disk to the IBM Flex System V7000 Storage Node.

7.2.3 Migration Wizard on the IBM Flex System V7000 Storage Node

Before starting the Migration, we want to make sure the remapped disks are available and visible on the IBM Flex System V7000 Storage Node GUI. Click **Pools** → **External Storage** and click **Detect MDisks** as shown in Figure 7-6 on page 271.

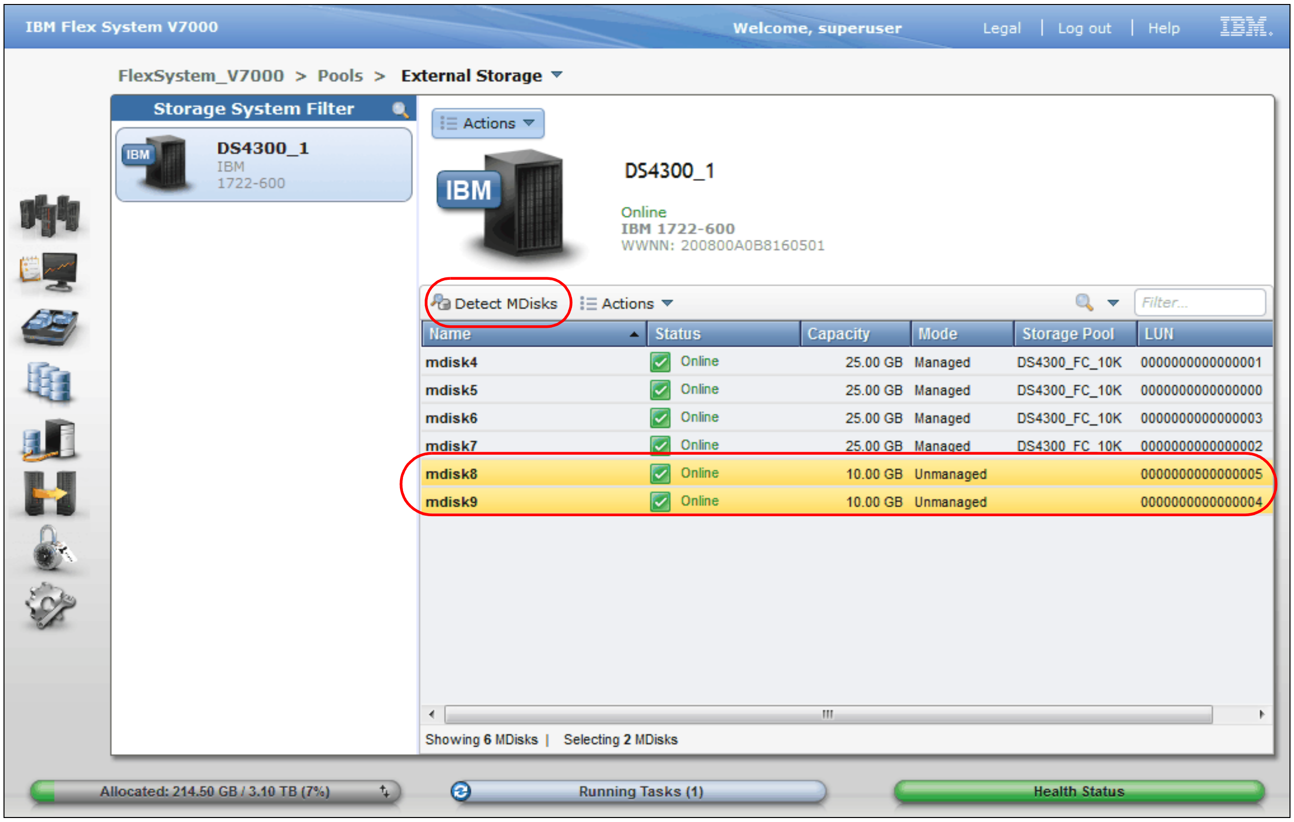


Figure 7-6 Remapped disks are visible in the External Storage section

There are two ways to access the menu options for starting a migration:

- Either Click **Home** → **Overview**, and then click **Migrate Storage** from the **Suggested Tasks** drop-down menu as shown in (Figure 7-7 on page 272).

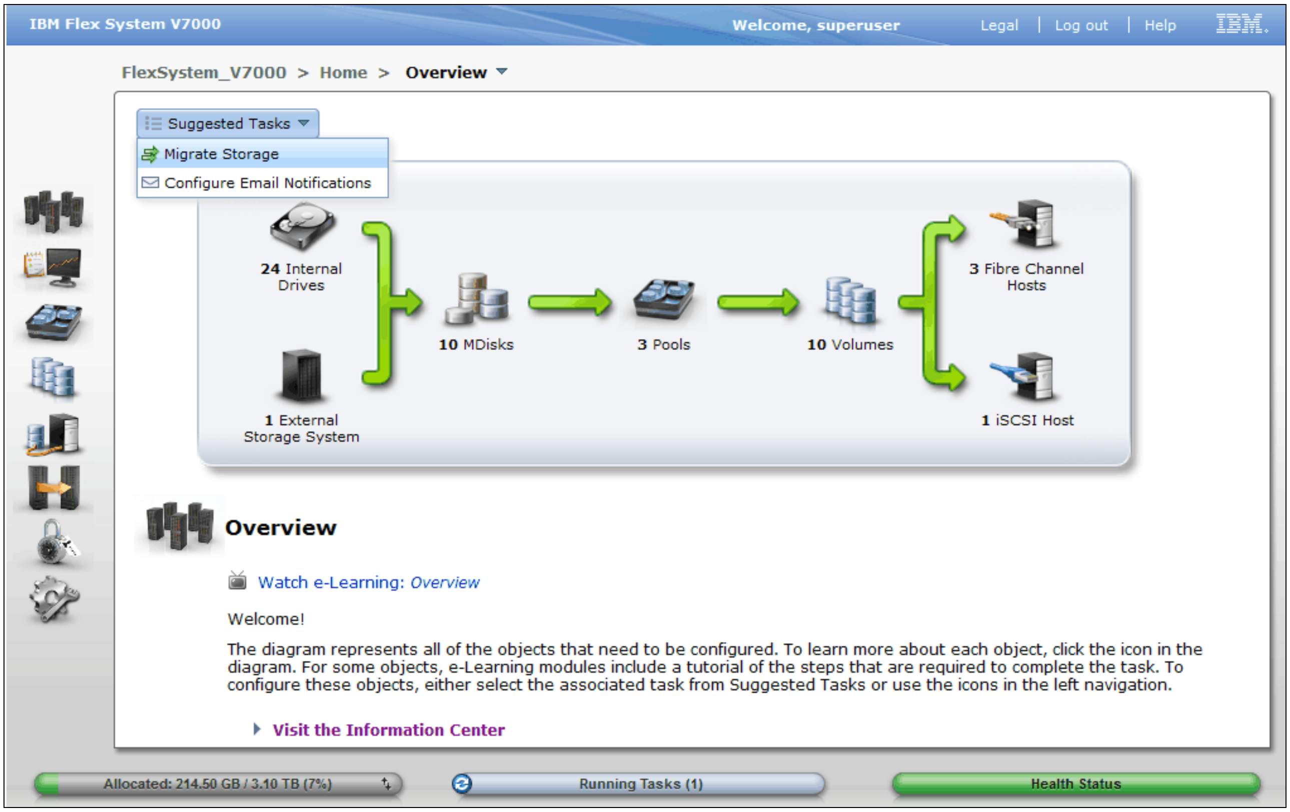


Figure 7-7 Migrate Storage from Overview Migrate Storage option

- Or Click the **Pools** icon and click **System Migration** as shown in Figure 7-8.

Figure 7-8 Migrate Storage from Pools menu System Migration option

Whichever method is chosen, the storage migration window opens (Figure 7-9 on page 274). Click **Start New Migration** to start the storage migration wizard.

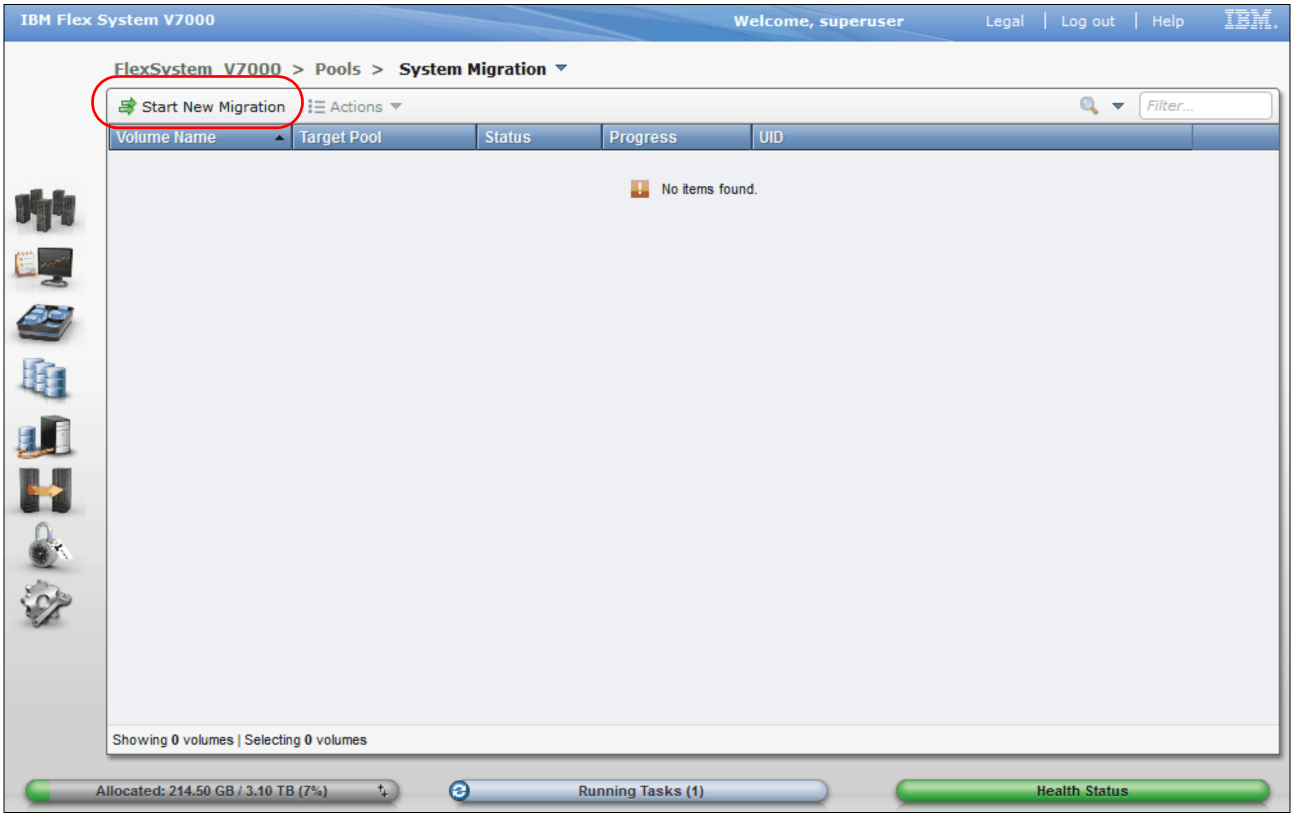


Figure 7-9 Storage Migration Wizard - Start New Migration

Using the IBM Flex System V7000 Storage Node storage migration wizard, you can easily migrate your existing data. Complete the following steps:

1. The Step 1 of the Storage Migration Wizard is to check the restrictions and prerequisites, and click **Next** as shown in Figure 7-10 on page 275.

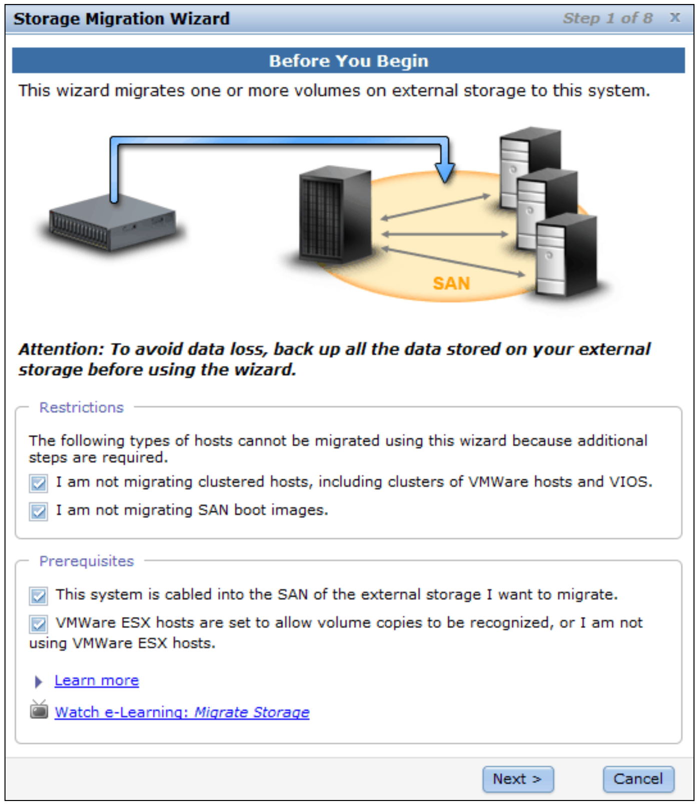


Figure 7-10 Storage Migration Wizard step 1- Before You Begin window

Avoiding data loss: To avoid any potential data loss, back up all the data stored on your external storage before using the wizard.

Figure 7-10 shows there are some restrictions and prerequisites to using the storage migration wizard:

– Restrictions:

- Do not use the storage migration wizard to migrate cluster hosts, including clusters of VMware hosts and VIOS.
- Do not use the storage migration wizard to migrate SAN Boot images.

If you have either of these two environments, you need to migrate them outside of this wizard. You can find more information about this topic at the IBM Flex System V7000 Storage Node Information Center at the following address:

http://publib.boulder.ibm.com/infocenter/flexsys/information/index.jsp?topic=%2Fcom.ibm.acc.common.nav.doc%2Fic-homepage.html

Author Comment: At this moment Then Flex System V7000 is not to find on the above link as the product is not yet GA.

The IBM Storwize V7000 Information Center also contain discussions about this topic:

http://pic.dhe.ibm.com/infocenter/storwize/ic/index.jsp

The VMware ESX Storage vMotion feature might be an alternative to migrating VMware clusters. For more information about this topic, see the following address:

http://www.vmware.com/products/vmotion/overview.html

– Prerequisites:
  - Make sure that the IBM Flex System V7000 Storage Node Fibre Channel ports have been connected to the SAN fabrics to which the external disk controller and hosts you want to migrate from are connected.
  - If you have VMware ESX hosts in the migration, make sure the VMware ESX hosts are set to allow volume copies to be recognized.

Click **Next** to go to the next step.

2. The Step 2 of the Storage Migration Wizard is to prepare the environment for migration, and click **Next** (Figure 7-11).

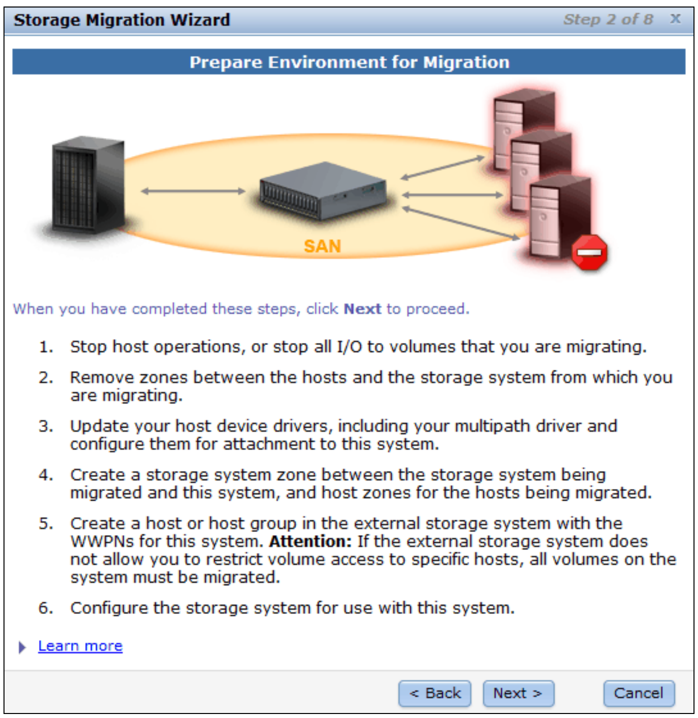


Figure 7-11 Storage Migration Wizard step 2 - Prepare the environment for migration

3. The Step 3 of the Storage Migration Wizard is to complete the mapping of external storage LUNS (Figure 7-12 on page 277). Make sure that you record the information mentioned in this step, as you need it for later steps. Click **Next**.

> **SCSI ID:** You might need to record the SCSI ID of the volume to which the host is mapped. Some operating systems do not support changing the SCSI ID during the migration.

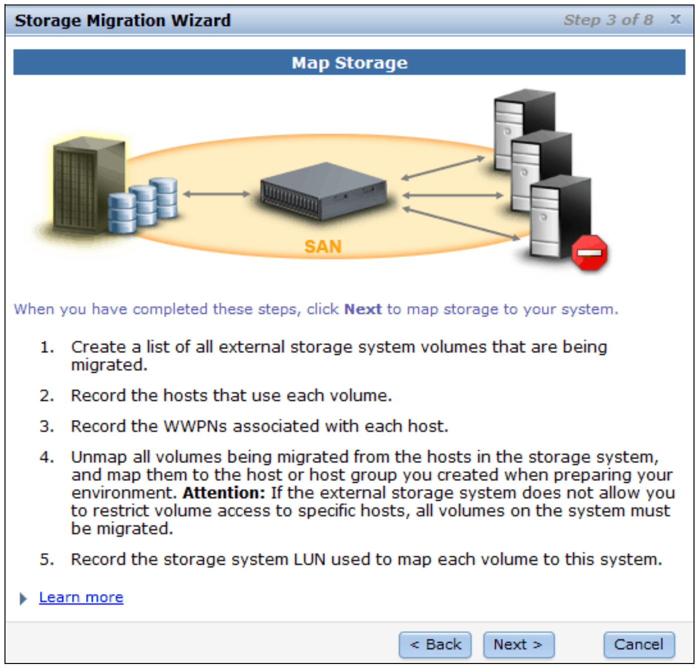


Figure 7-12 Storage Migration Wizard step 3 - Map Storage

After you click **Next,** the IBM Flex System V7000 Storage Node starts to discover external devices (if you have correctly zoned the external storage systems with the IBM Flex System V7000 Storage Node and mapped the LUNs). When the discovery completes, IBM Flex System V7000 Storage Node shows the MDisks found.

4. The Storage Migration Wizard performs MDisk scan automatically and any new or unused MDisks appear on the list. Choose the MDisks you want to migrate in Step 4 and click **Next** (Figure 7-13).

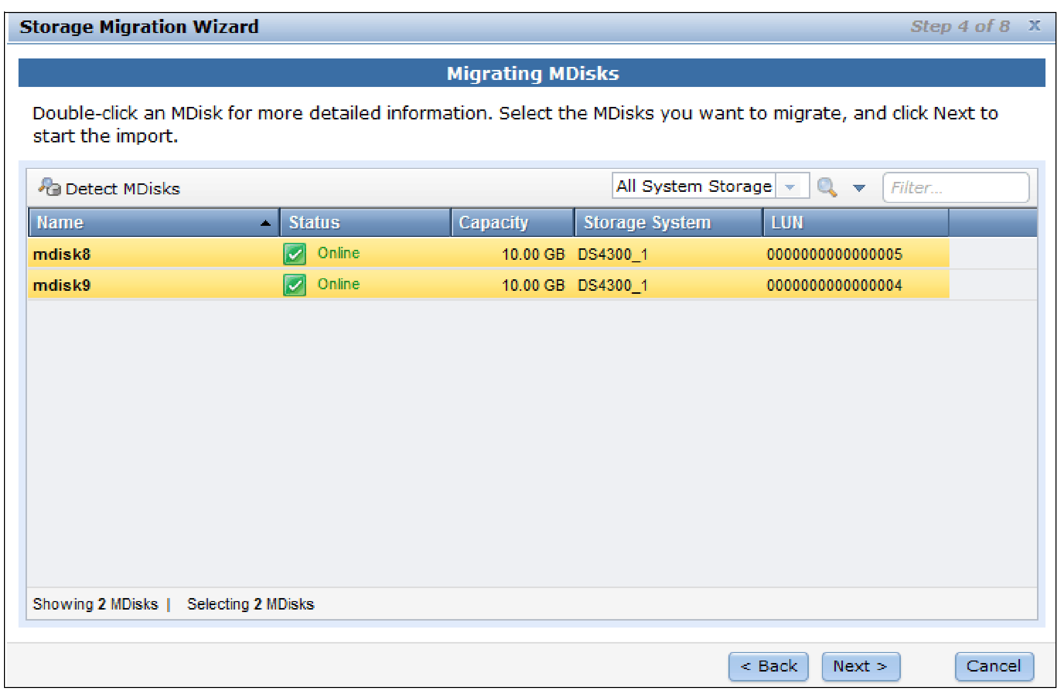


Figure 7-13 Storage Migration Wizard step 4 - Select MDisk

If the MDisks that need migrating are in the list, select them and click **Next**. The IBM Flex System V7000 Storage Node starts to import the MDisks that you have chosen. If the MDisks that need migrating are not in the list, you might need to check your zone configuration and LUN mapping, and click **Detect MDisks** to run the IBM Flex System V7000 Storage Node discovery procedure again.

You can select one or more MDisks as required, and detailed information about the MDisk can be shown by double-clicking it.

In Figure 7-13 on page 277, two LUNs are discovered as MDisks that are candidates for migration. In your particular situation, you might need to reference the information you recorded earlier to identify these MDisks. In our example, the MDisks have been selected to go forward to the next step.

When you click **Next** in this step, IBM Flex System V7000 Storage Node completes the importation of the MDisks with the host's data and a storage pool has been created. The MDisks are added to the pool and image mode volumes (with the same size as the MDisks) are created, which are ready for mapping back to the original hosts.

The Storage Migration Wizard creates a special migration pool that contains the copy source and the copy destination volumes. When the Storage Migration has finished and it has been Finalized by the administrator then the copy source MDisk returns to an unmanaged MDdisk and the copy destination is now part of the volume that was selected during the Storage Migration Wizard.

Note: If an unmanaged MDisk is manually added to a pool from the MDisks by Pool Menu then this MDisk becomes managed and data is striped across it.

5. Configure the host that needs to access the data after the migration, or create new hosts as needed in Step 5, and click **Next** (Figure 7-14).

Figure 7-14 Storage Migration Wizard step 5 - Configure Hosts

Before you configure any hosts, make sure that the appropriate drivers have been installed on the host and that the host zones have been zoned correctly.

If the host that needs to access the data on the volume after the migration is complete is in the list, select the host and click **Next**.

If the host has not been created on the IBM Flex System V7000 Storage Node storage system, click **New Host** to create it and follow the instructions for creating the new host. In our example we assume that the IBM Flex System V7000 Storage Node already has been configured with the correct host and corresponding World Wide Names or iSCSI addresses. After the host has been created, you can find it in the host list.

6. Click **Map to Host** for making the disks available to the host during migration as shown in Figure 7-15 on page 280.

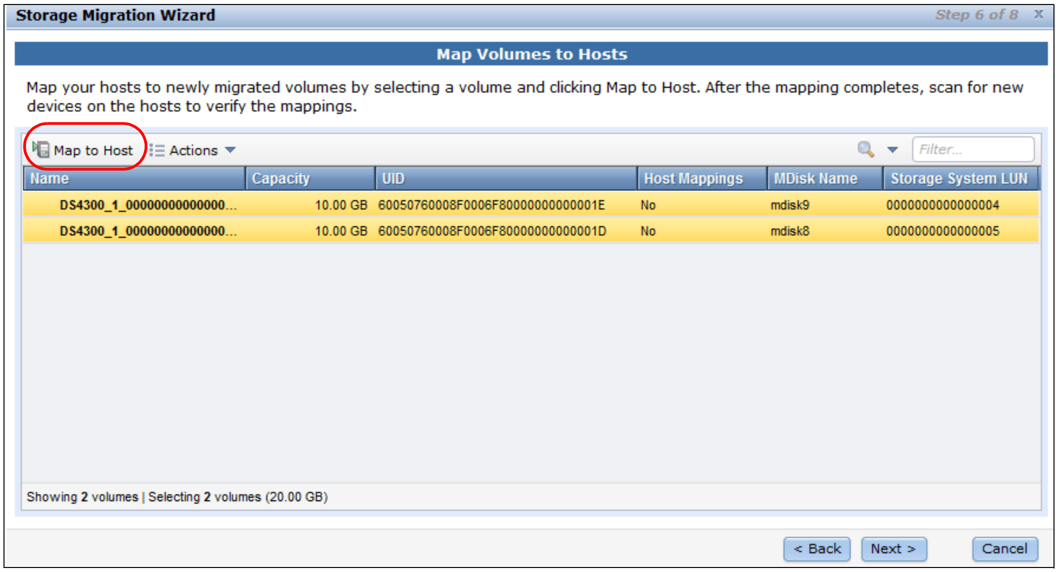


Figure 7-15 Storage Migration Wizard step 6 - Map to host

7. Select the host you want to use for mapping the volume. Click **Next** to continue with the migration wizard as shown in Figure 7-16.

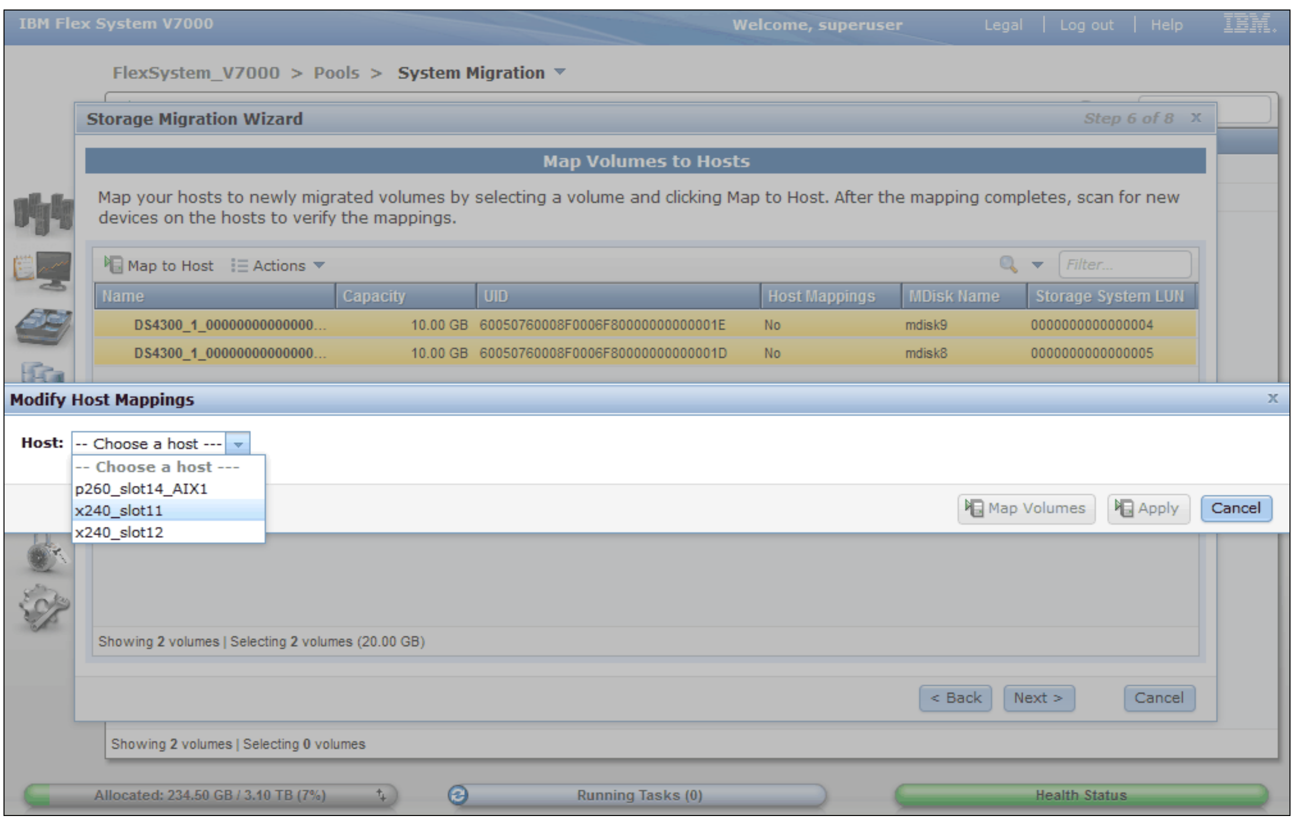


Figure 7-16 Storage Migration Wizard - Map disk to Windows 2008 host

8. Map the newly migrated volume to the host, and when the mapping is complete, click **Apply** or **Map Volumes** as shown in Figure 7-17 on page 281.

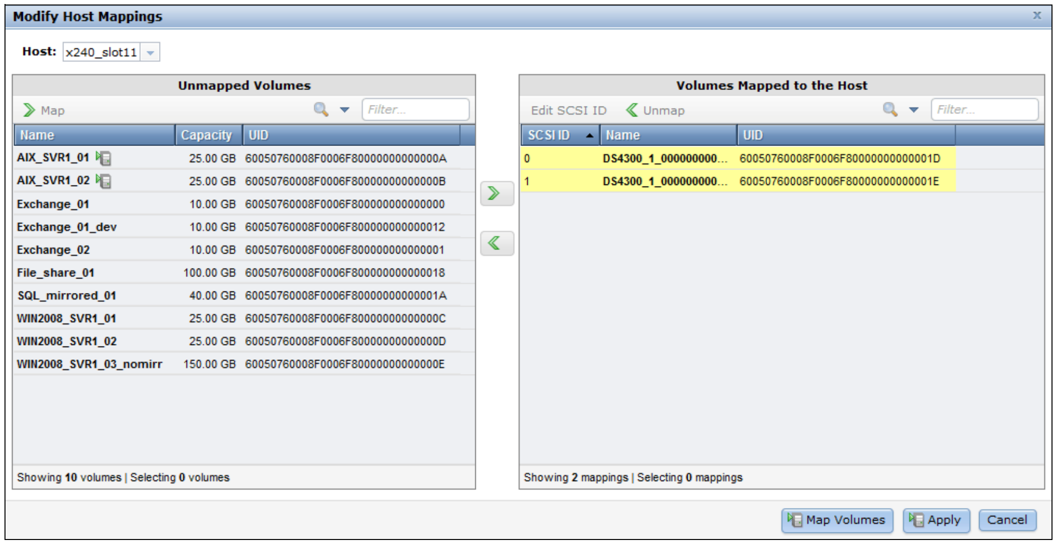


Figure 7-17 Storage Migration Wizard s- Apply Mapping to Windows 2008 host

In Step 4 on page 277 of the migration wizard, the volumes from the imported MDisks that need to be migrated have been listed. The names of the volumes have been assigned automatically by the IBM Flex System V7000 Storage Node storage system. You can change the names to any words that would be meaningful to you by selecting the volume and clicking **Rename** in the **Actions** drop-down menu.

Names: The names must begin with a letter. The name can be a maximum of 63 characters. Valid characters are uppercase letters (A-Z), lowercase letters (a-z), digits (0 - 9), underscore (_), period (.), hyphen (-), and space. The names must not begin or end with a space.

To map the volumes to the hosts, select the volumes and click **Map to Host**. A window opens with a drop-down list of the hosts. At this time, the IBM Flex System V7000 Storage Node lets you choose which host to which you need to map the volumes. Choose the correct host and click **Next**.

Volume mapping: As a best practice, map the volume to the host with the same SCSI ID before the migration, which you should have recorded in Step 3 on page 276.

When the mapping completes, you discover that the Host Mappings column of the volumes changed from No to Yes. A scan could be performed to discover the new devices on the host for verification. Click **Next** to go to the next step of the storage migration wizard.

Figure 7-18 on page 282 shows Step 6 of the storage migration wizard with host mappings modified.

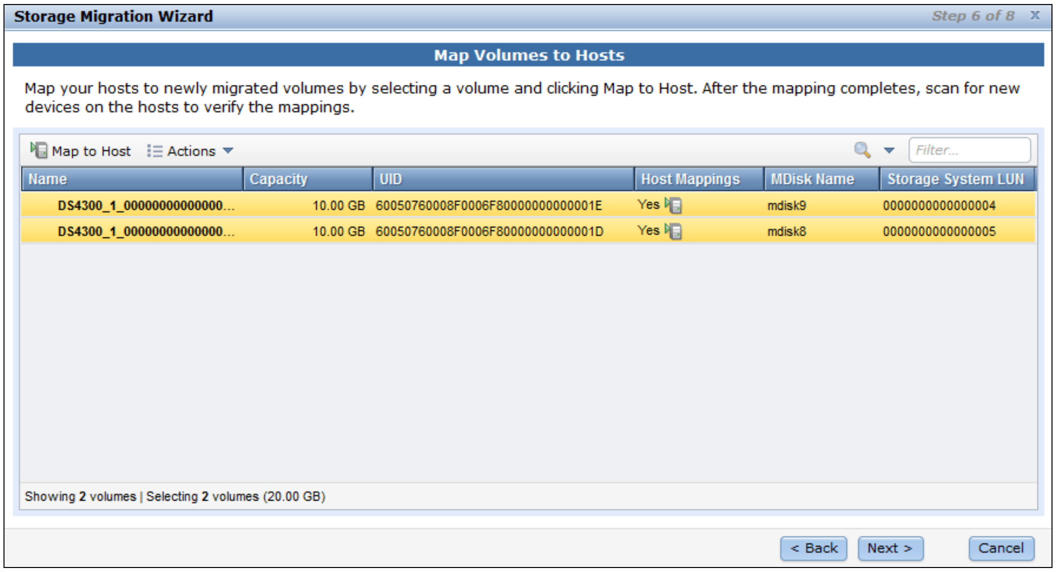


Figure 7-18 Storage Migration Wizard - Completion of the host mappings

9. Select the destination storage pool for data migration in Step 7 and click **Next** (Figure 7-19).

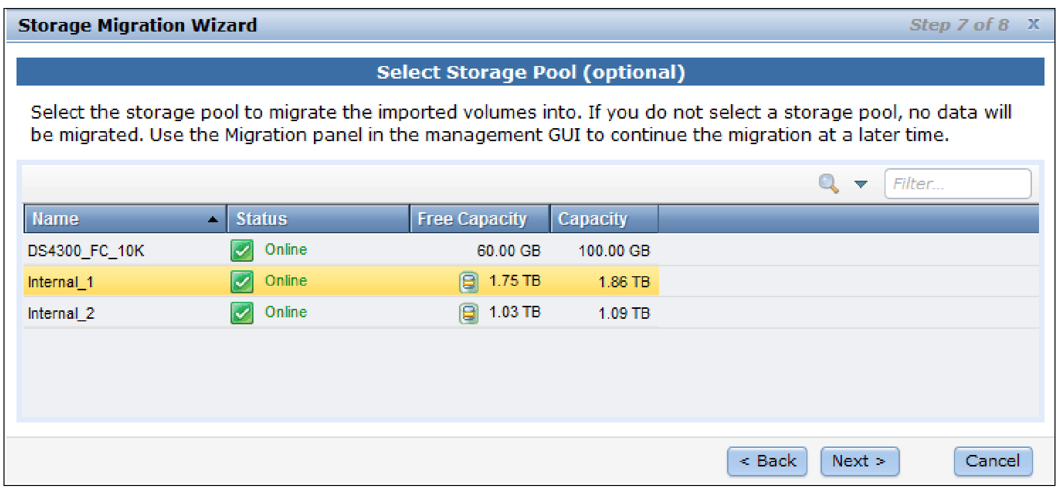


Figure 7-19 Storage Migration Wizard step 7 - select Storage Pool

The destination storage pool of the data migration could be an external storage pool or an internal storage pool. Make sure that there is enough space in the storage pools.

After you click **Next**, the migration begins. The migration runs in the background and results in a copy of the data being placed on the MDisks in the storage pool selected. The process uses the volume mirroring function included with the IBM Flex System V7000 Storage Node, and when complete, the volumes have pointers to both the new copy on the storage pool selected and on the original external storage system.

10. Click **Finish** to end the storage migration wizard in Step 8 as shown in Figure 7-20 on page 283.

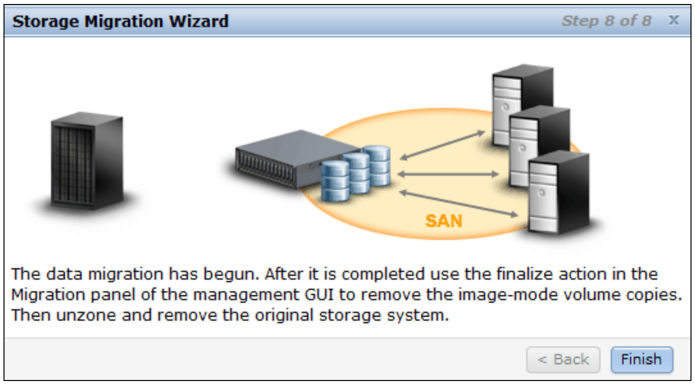


Figure 7-20 Storage Migration Wizard step 8 - Migration has begun

When returning to the **Pools** → **System Migration** window the migration progress is showing. Also at the status indicators Running Tasks show that volume synchronization is in progress as shown in Figure 7-21

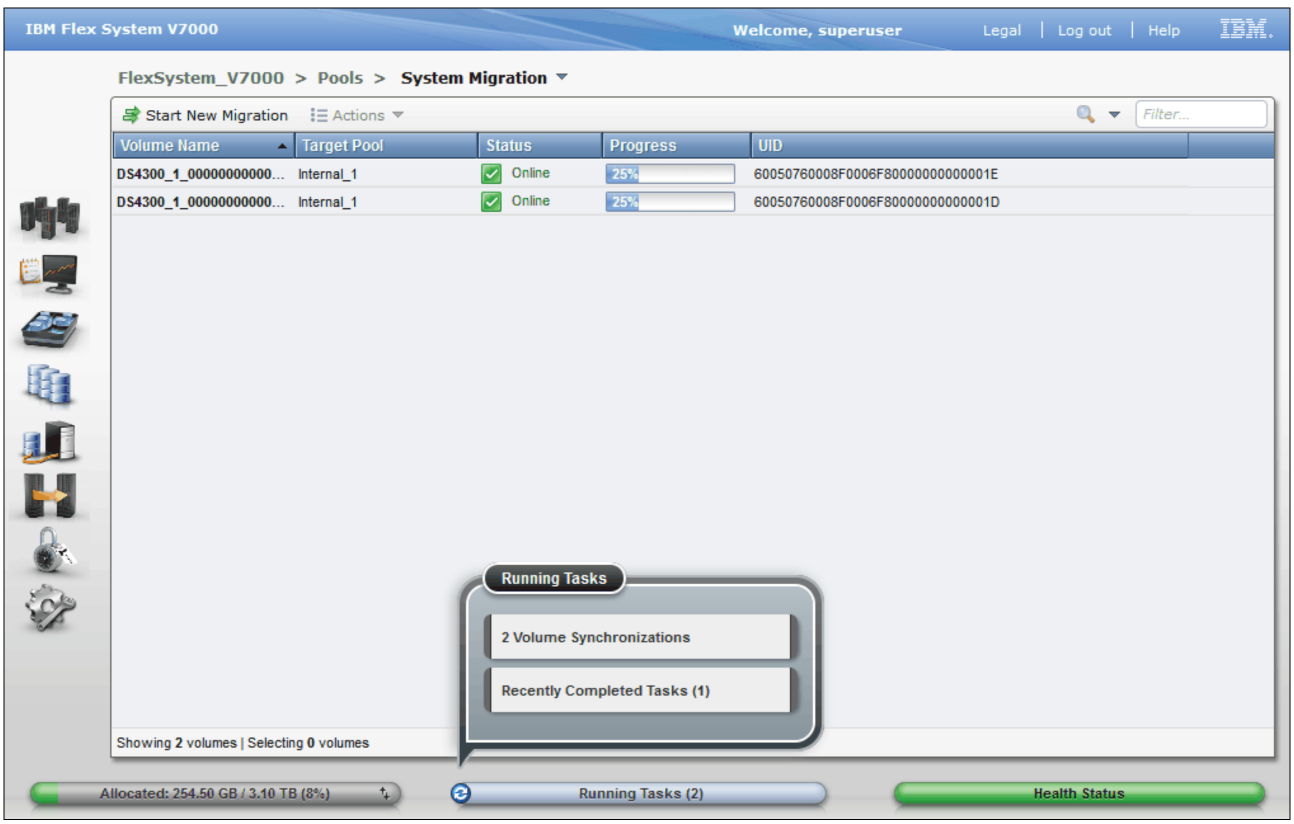


Figure 7-21 Migration is in progress

7.2.4 Checking that the Windows server can see its disks

Immediately after the Storage Migration Wizard has finished as shown above, the disks are available to the Microsoft Windows 2008 Server and you may reboot this or rescan its disks to get access to its data again.

Data may still be copying to the destination storage pool. This process may take hours or even days depending of the volume size and system load.

Figure 7-22 shows that the Microsoft Windows 2008 Server now again has access to the migrated disks. Data is still being copied from the IBM DS4300 disk system to the IBM Flex System V7000 Storage Node as a background process.

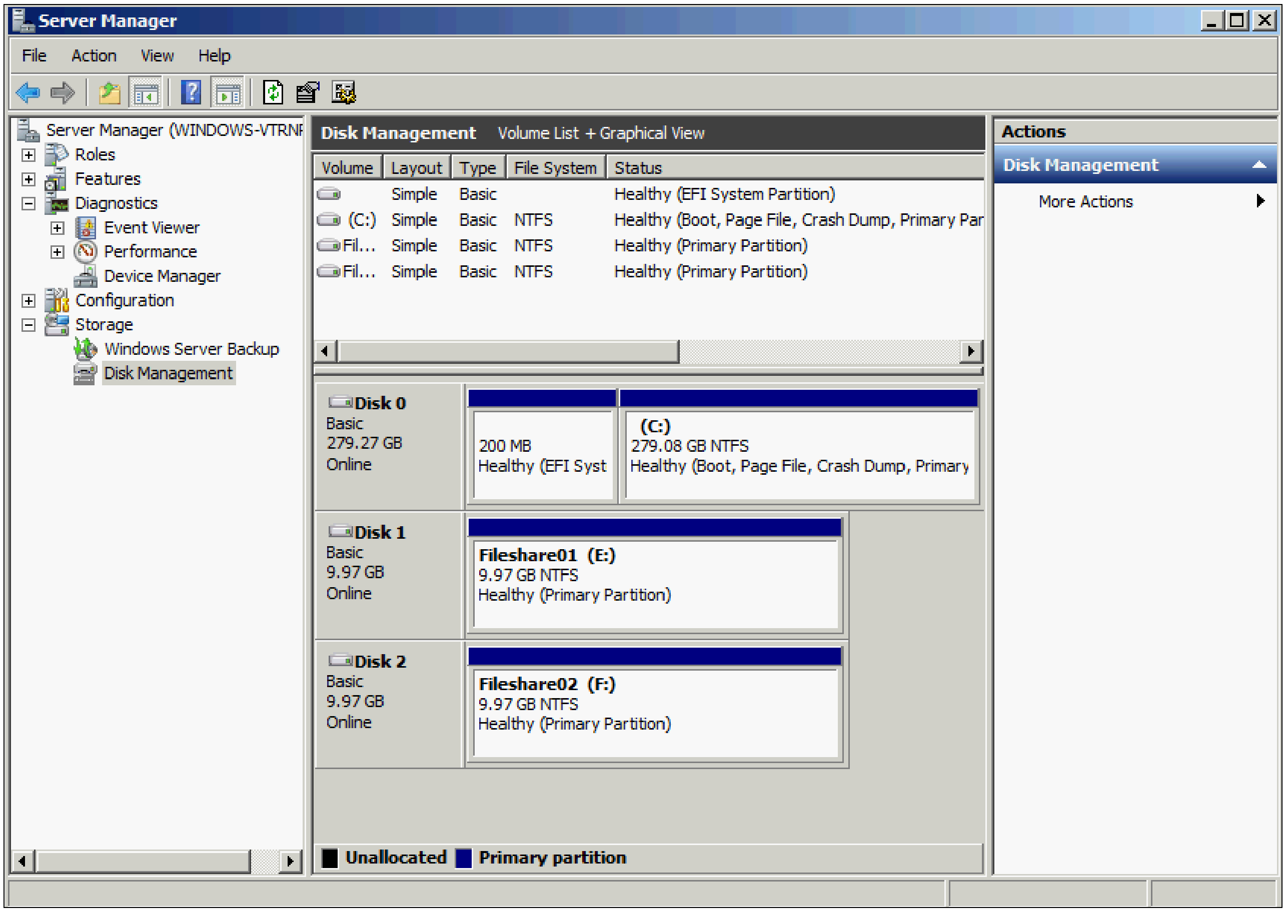


Figure 7-22 Windows server with disk from IBM Flex System V7000 Storage Node

While data is being copied the IBM Flex System V7000 Storage Node creates a mirror copy between the source disk at the IBM DS4300 and the IBM Flex System V7000 Storage Node. When the migration process finishes all you need to do is to Finalize the migration. By Finalizing the migrating, the link between the two mirror copies is broken, and the source disks at the IBM DS4300 return as unmanaged MDisks to the IBM Flex System V7000 Storage Node (shown in Figure 7-27 on page 288).

Figure 7-23 on page 285 shows that the disk is in a migration pool which can be reviewed by clicking **Pools** → **Volumes by Pool.** It remains here until the migration process is Finalized.

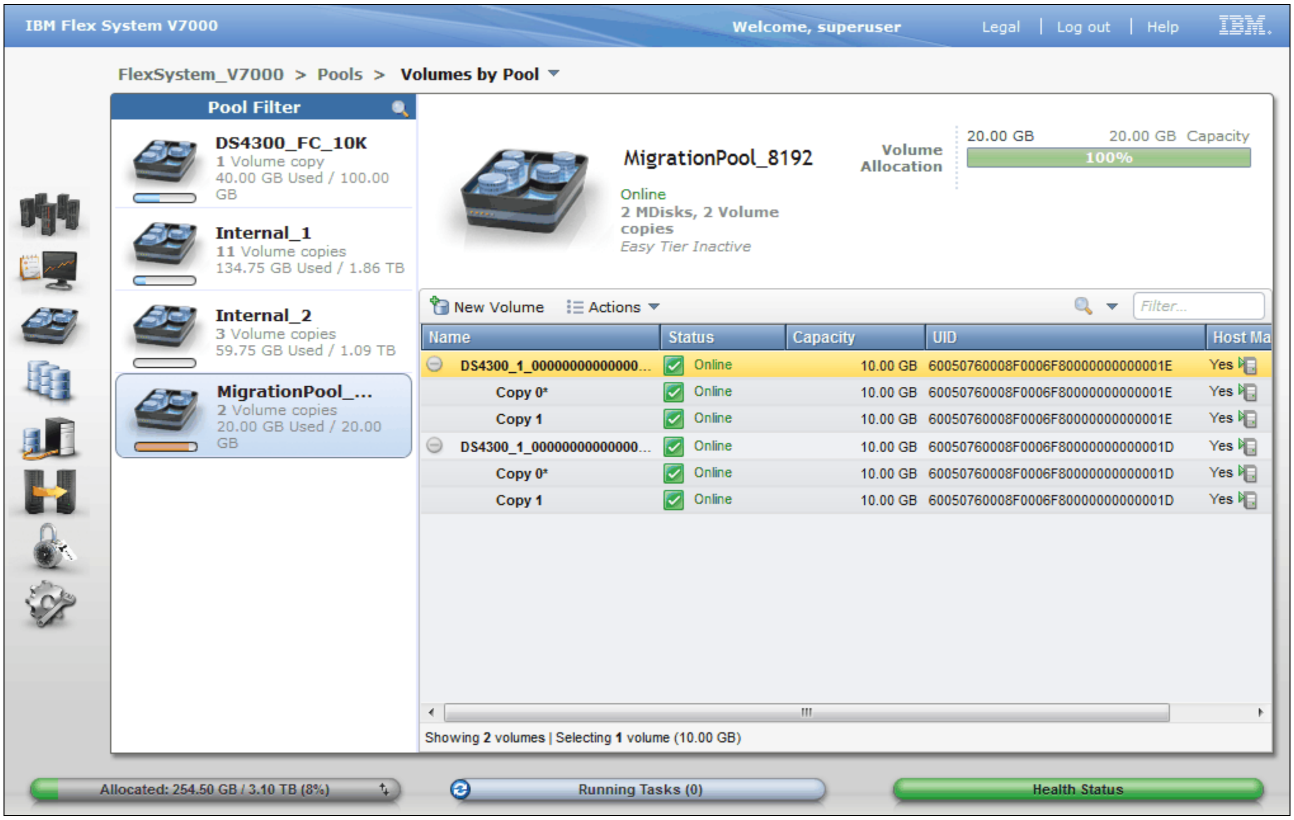


Figure 7-23 After Migration - Disk is now in a Migration pool

The end of the storage migration wizard is not the end of the data migration. The data migration has begun, and after clicking **Finish** in step 10 on page 282, you can find the migration progress in the migration window. You can also find the target storage pool to which your volumes are being migrated, along with the status of the volumes.

Figure 7-24 on page 286 shows the data migration progress in the migration window at the time where the copy process reached 100% which can be reviewed by clicking **Pools** → **System Migration**.

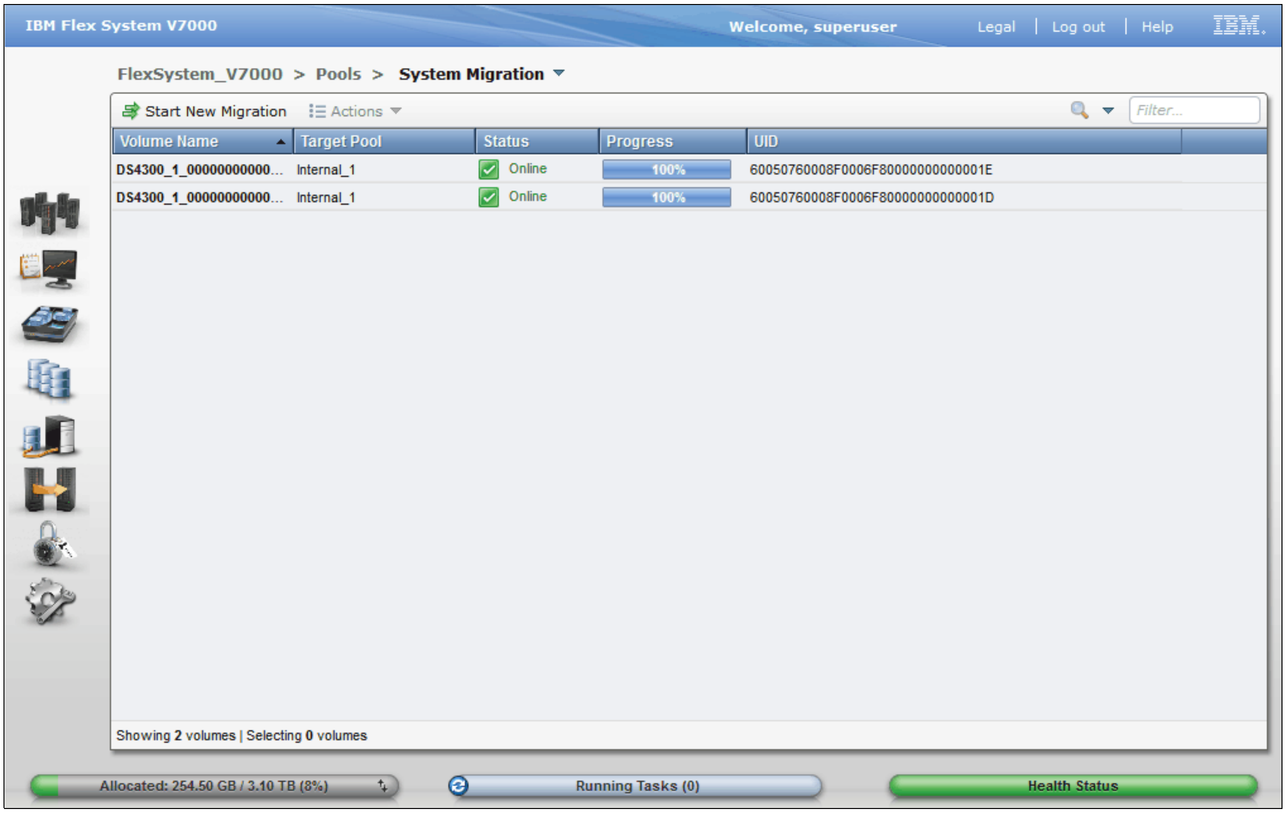


Figure 7-24 Migration progress in the migration window

7.2.5 Finalizing the migration

When the migration progress reaches 100%, select one or both volumes and click **Finalize** in the **Actions** drop-down menu in the migration window (Figure 7-25 on page 287).

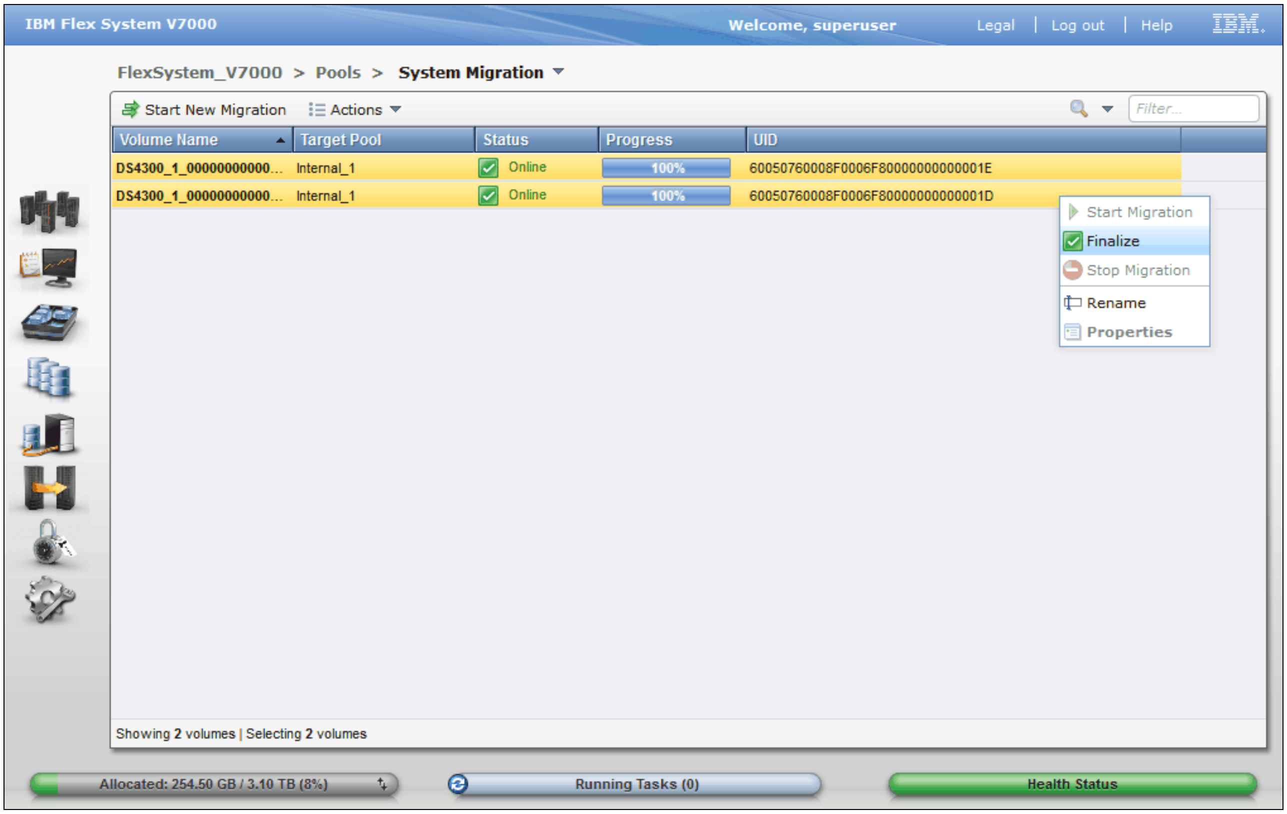


Figure 7-25 Storage Migration Wizard - Finalize migration step 1

After you select **Finalize** in the **Action** menu, the IBM Flex System V7000 Storage Node needs you to confirm the migration of the volumes. Verify the volume name and the number of migrations and click **OK**. Figure 7-26 shows the Finalize Volume Migration window.

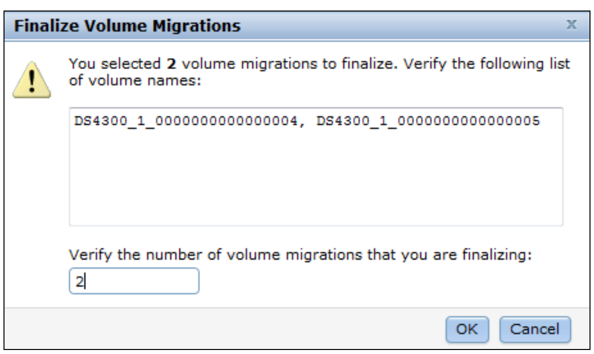


Figure 7-26 Storage Migration Wizard - Finalize migration step 2

When the finalization completes, the data migration to the IBM Flex System V7000 Storage Node is done.

The image mode copy of the volumes on the original external storage system is deleted and the associated MDisks from the storage pool are removed and the status of those MDisks then becomes unmanaged. They are now available to the system as unmanaged disks as shown in Figure 7-27 on page 288

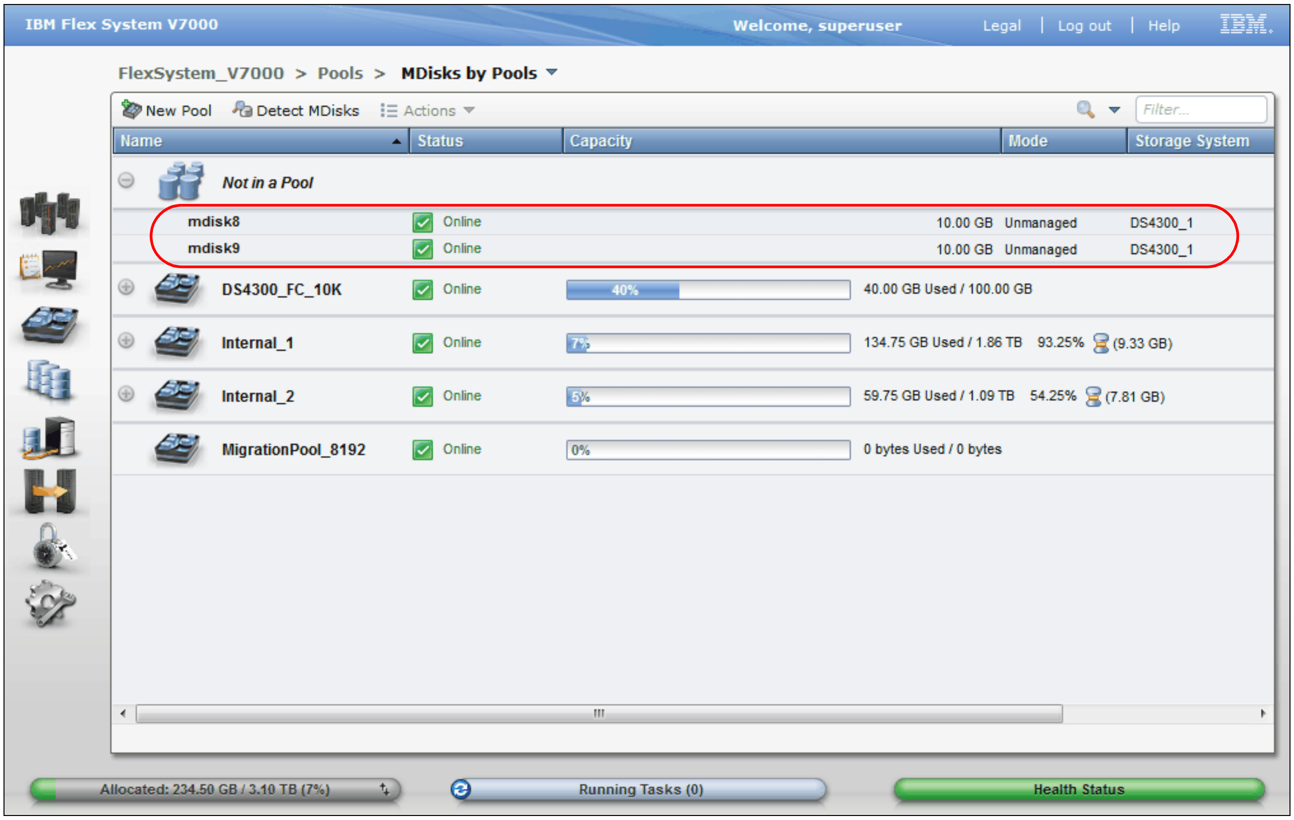


Figure 7-27 Disks are returned as unmanaged MDisks

You can now un-zone and remove the older storage system from the IBM Flex System V7000 Storage Node.

For more information about the IBM Flex System V7000 Storage Node migration function, refer to Chapter 10, “Volume Mirroring and Migration” on page 423.

After the migration is Finalized, the Migration Pool used for the migration remain as a Storage Pool on the system and is empty. It can be deleted at this time or it can simply be left unused.

As the volume name of the migrated disk is automatically determined by the Storage Migration Process, and if you did not rename the disk during the Wizard this is now a good time to rename the volume to a name that reflects what it is used for. For an example of how to rename a volume refer to 7.2.7, “Renaming the volume” on page 291

7.2.6 Mapping to host performed after the Migration has begun

If you did not map the disk to the host during the migration Wizard, mapping can be performed now. Follow these steps to map the disk:

- Figure 7-28 on page 289 shows how to navigate to **Pools** → **Volumes By Pool** from where we locate the volume being migrated - right click the volume and click **Map to Host**. The same can be done by selecting the volume and from the **Actions** menu select **Map to Host**.

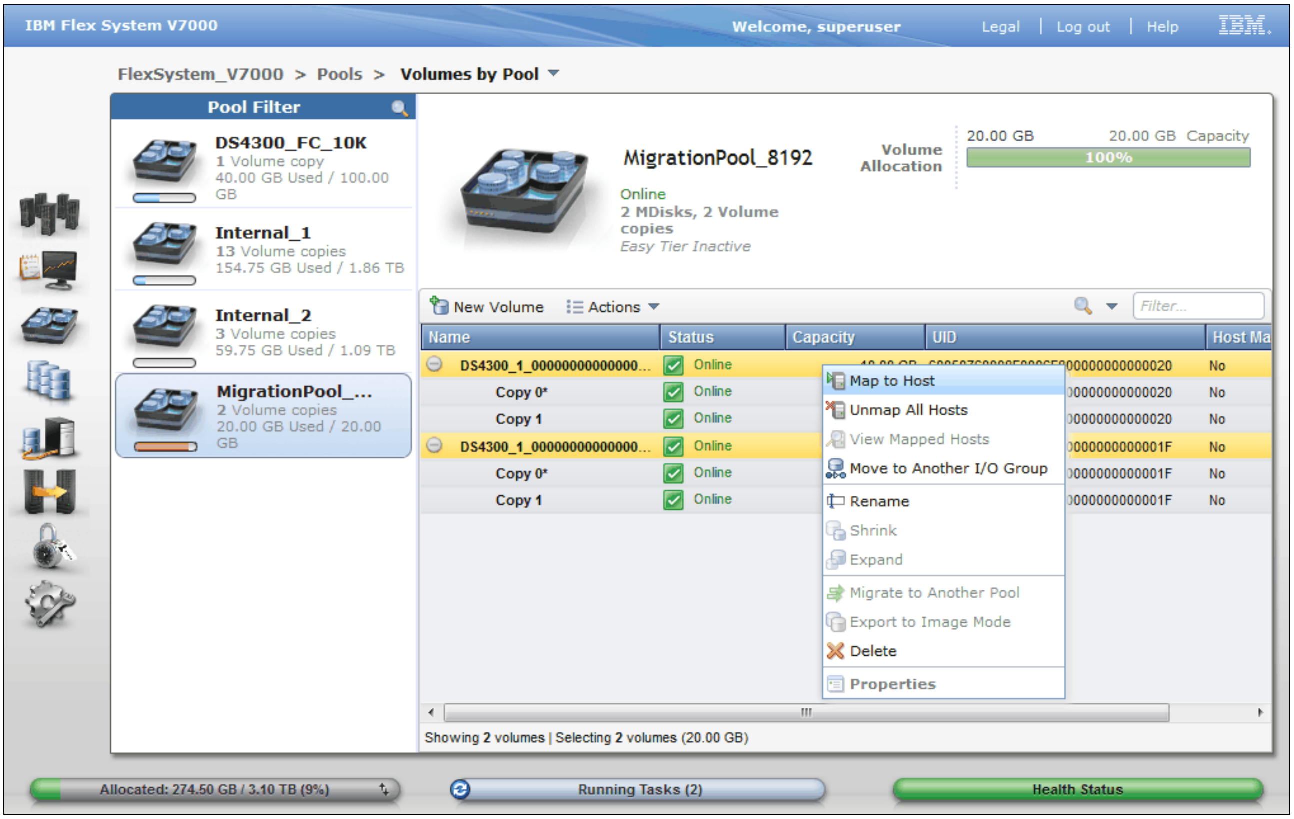


Figure 7-28 After Migration start - Map to host

- Figure 7-29 on page 290 shows the next step of the mapping process. The wizard shows all the hosts currently configures on the IBM Flex System V7000 Storage Node and lets us select which host we want to map the volume.

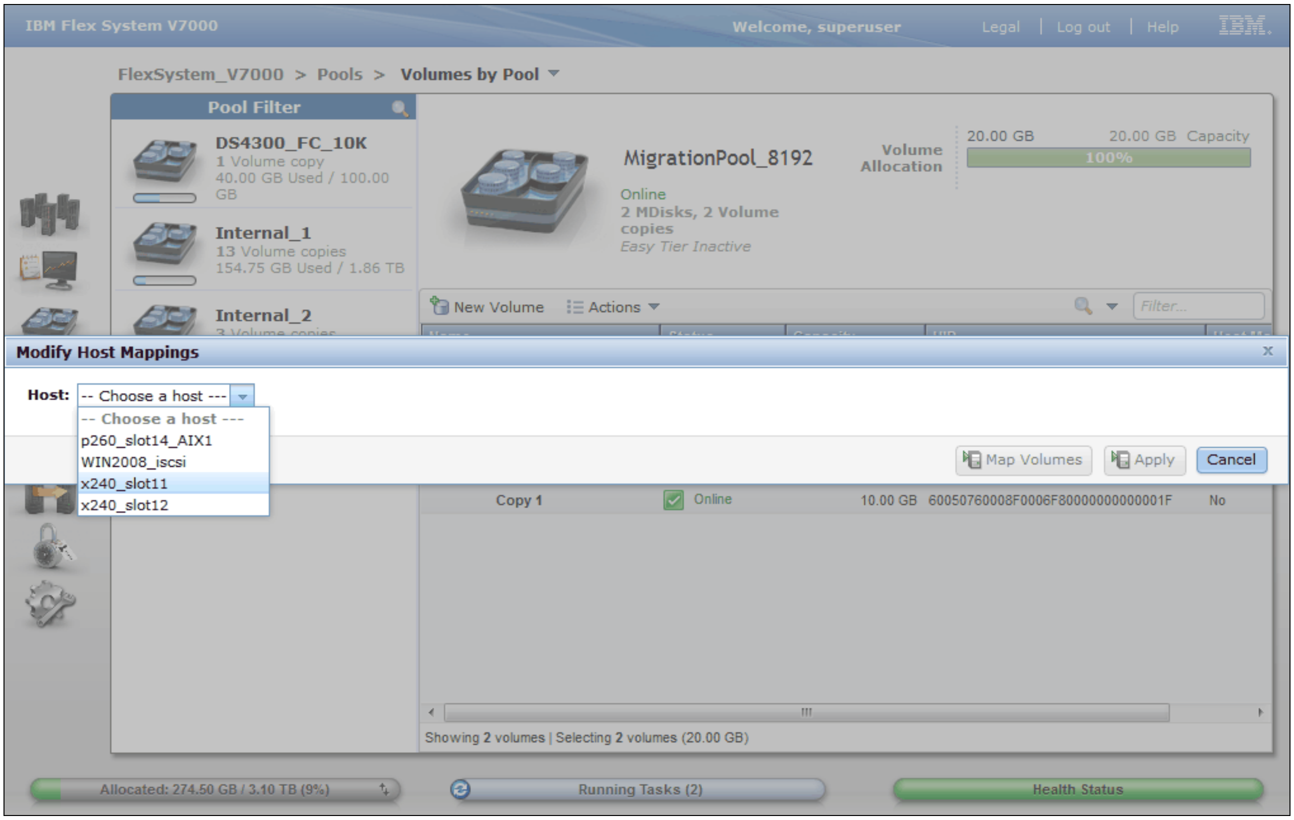


Figure 7-29 Modify Host Mappings - Choose a host

When the host is selected you can click either **Map Volumes** or **Apply** whichever will continue to the next step of the Map to Host Wizard.

- Figure 7-30 shows all unmapped volumes that can be mapped to the host. The volume we right-clicked in the first step now appears in the right-side of the window and is ready to be mapped.

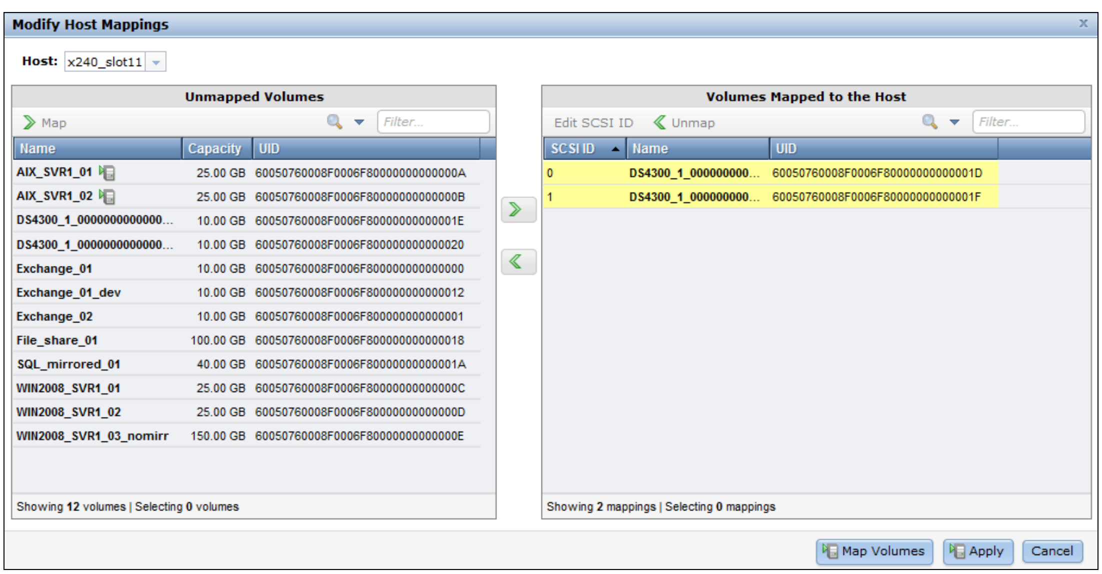


Figure 7-30 Modify Host Mappings

You can change the SCSI ID of the new mappings through the **Edit SCSI ID** button or by double clicking the **SCSI ID** Field. Clicking either **Map Volumes** or **Apply** maps the volume to the host.

Click **Close** on the CLI commands window to complete the mapping.

7.2.7 Renaming the volume

The migration wizard automatically creates new names for the volumes being migrated and since the names relate from its previous name the new name may no longer be appropriate or indicative to where the volume resides.

Figure 7-31 shows how we review the volume name that has automatically been given to the migrated volume. This can be done by selecting **Volumes** → **Volumes by Host**. More volumes can be selected and renamed at the same time.

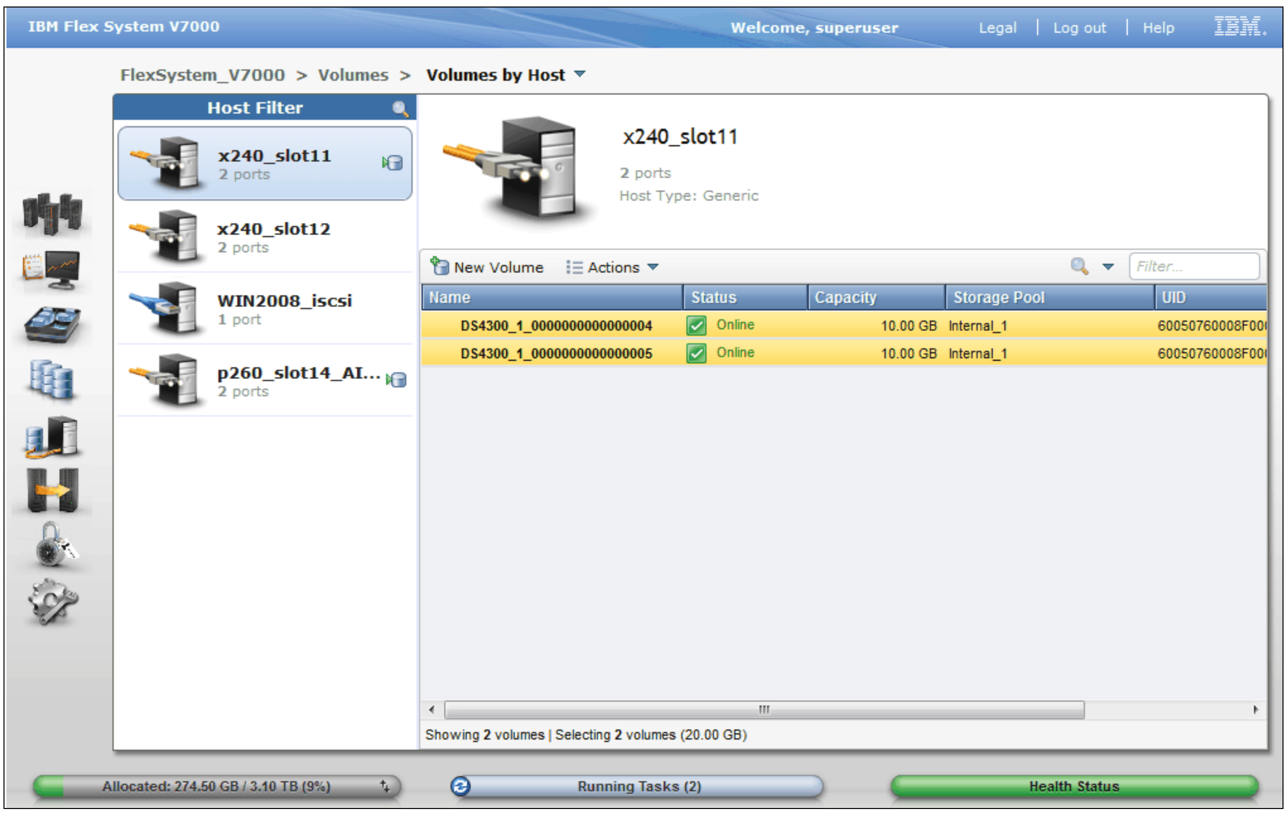


Figure 7-31 After Migration - Volume Name before rename

After choosing the host, the Modify Mappings window opens. On the right, you can find your newly mapped volumes highlighted in yellow.

Figure 7-32 on page 292 shows how we rename the volume. This is done by selecting **Volumes** → **Volumes by Host** → **Actions** → **Rename.**

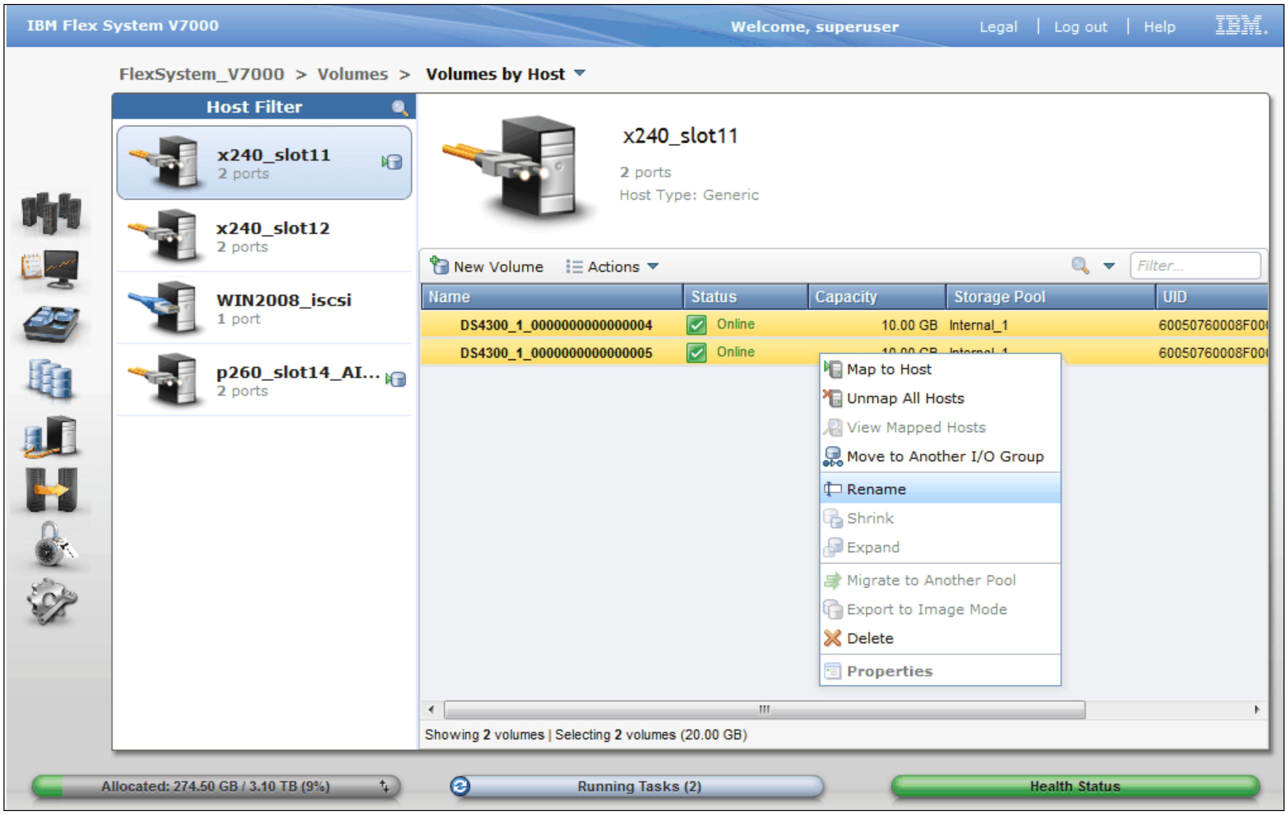


Figure 7-32 After Migration - Rename volume step 1

Figure 7-33 show how we give the volume a name that makes sense according to which function the volume has.

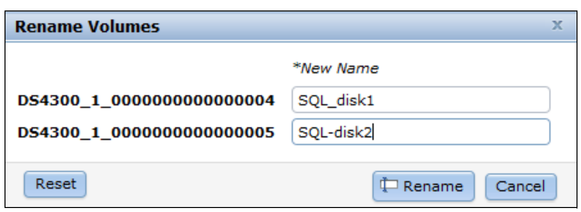


Figure 7-33 After Migration - Rename Volume step 2

Figure 7-34 on page 293 shows that the volume has been renamed.

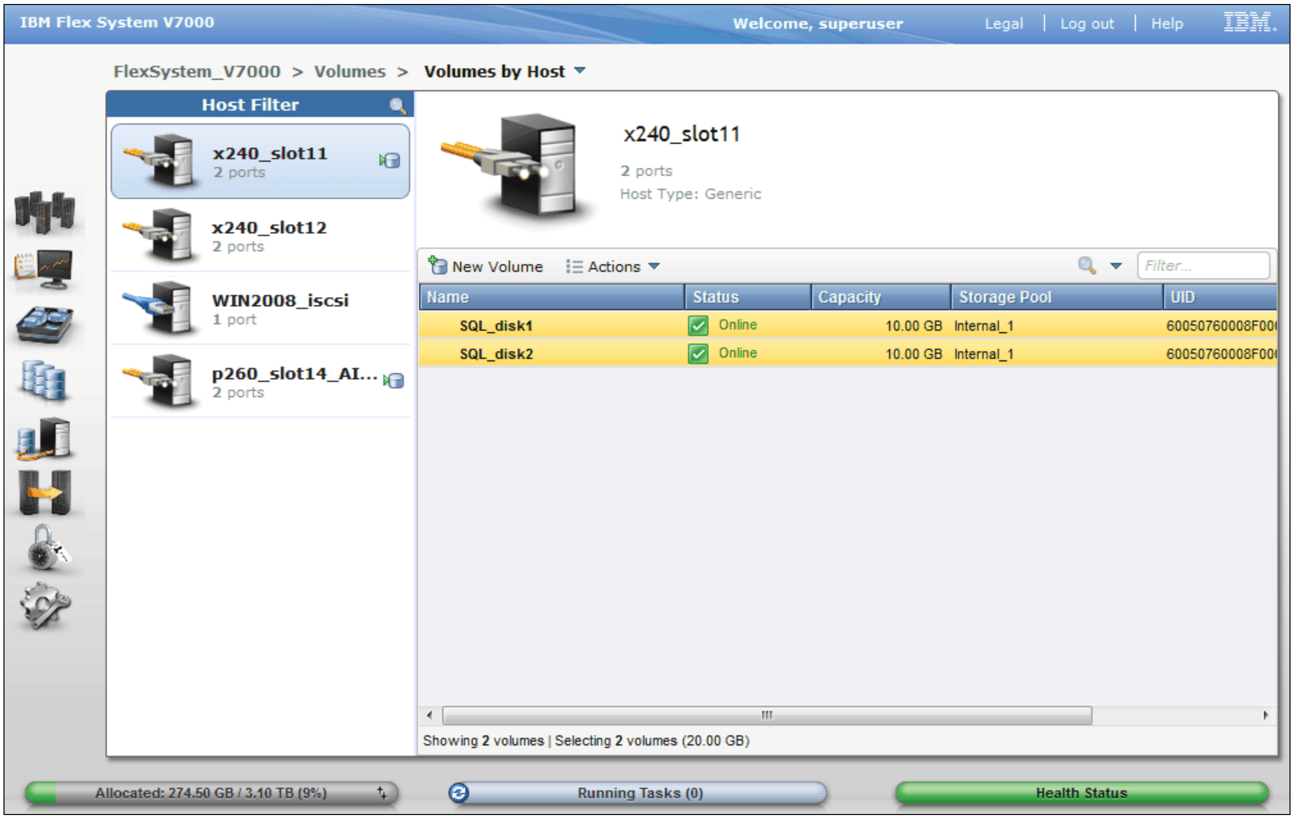


Figure 7-34 After Migration - Volume Name after rename

8

Storage pools

This chapter explains how IBM Flex System V7000 Storage Node manages physical storage resources. All storage resources under IBM Flex System V7000 Storage Node control are managed using *Storage pools*. Storage pools and the volumes you build on them make it easy to dynamically allocate resources, manipulating these and hence maximize productivity and reduce costs.

8.1 Working with internal drives

This section describes how to configure the internal storage disk drives using different RAID levels and different optimization strategies.

In this section, we start with the environment shown in Figure 8-1. All the internal drives are unconfigured. Currently, the existing MDisks come from external storage, and example storage pools, volumes, and hosts have been created for use.

You can learn how to manage MDisks and storage pools in this chapter. In Chapter 7, “Migration wizard” on page 265, we demonstrate how to discover and import LUNs from external disk array controllers and how to migrate data out of these and into IBM Flex System V7000 Storage Node. You can also learn about, how to create volumes and hosts and how to map volumes to hosts in Chapter 6, “Basic Volume and Host Configuration” on page 241.

The IBM Flex System V7000 Storage Node storage system provides an individual Internal Storage window for managing all internal drives. The Internal Storage window can be accessed by opening the Overview window, clicking the **Internal Drives** function icon and then clicking **Pools**. Figure 8-1 shows you how to access the Internal window from the Getting Started window.

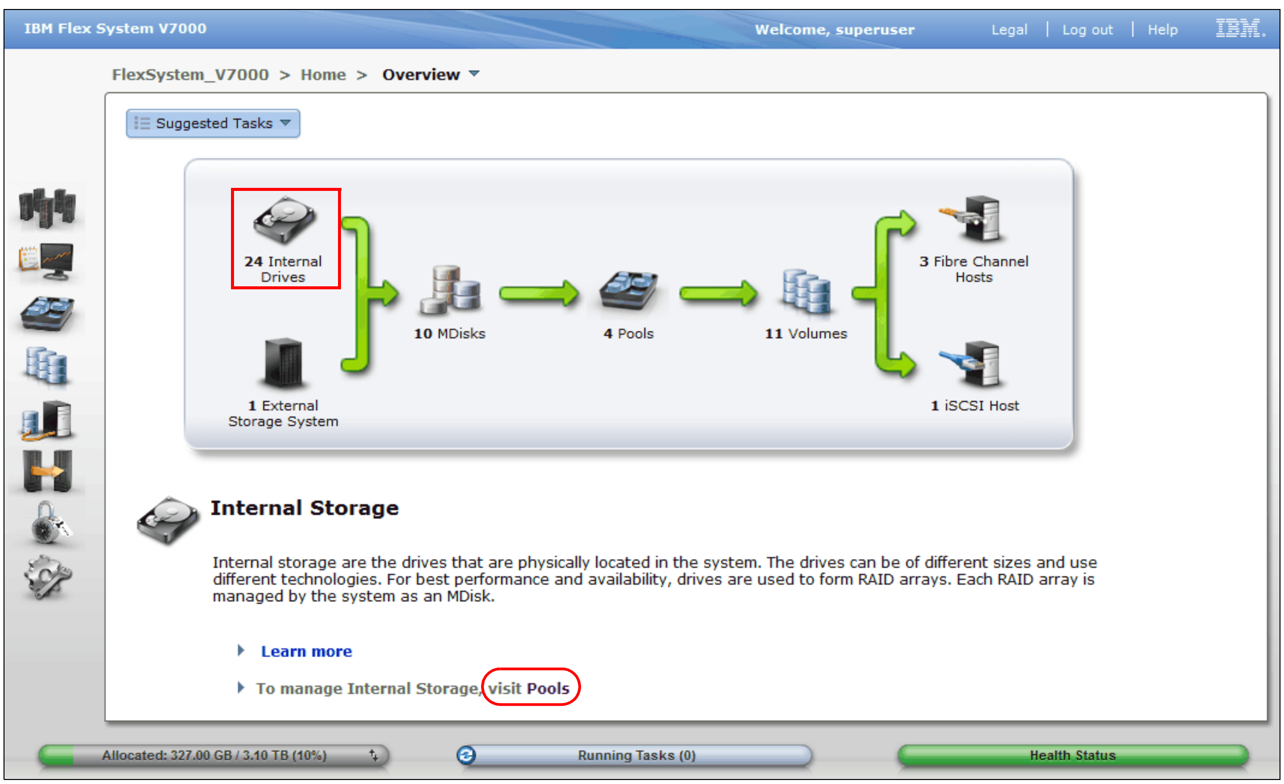


Figure 8-1 Access the Internal Storage window from the Overview screen

An alternative way to access the Internal Storage window is by clicking the **Pools** icon on the left side of the window (Figure 8-2).

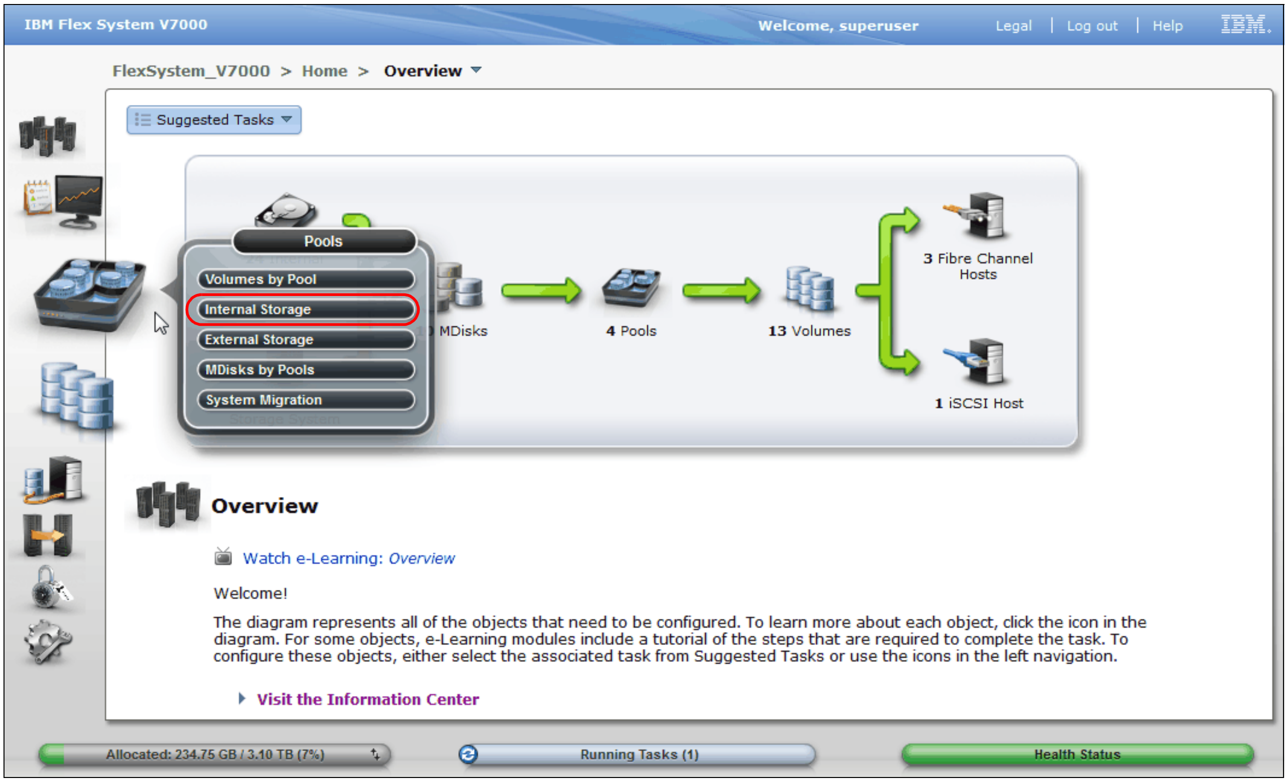


Figure 8-2 Access the Internal Storage window using the Pools icon

The Internal Storage window (Figure 8-3) gives an overview of the internal drives installed in the IBM Flex System V7000 Storage Node storage system. Selecting **All Internal** in the **Drive Class Filter** shows all the drives installed in the managed system, including attached expansion enclosures. Alternatively, you can filter the drives by their type or class, for example, you can choose to show only SAS, NL SAS, or SSD drives.

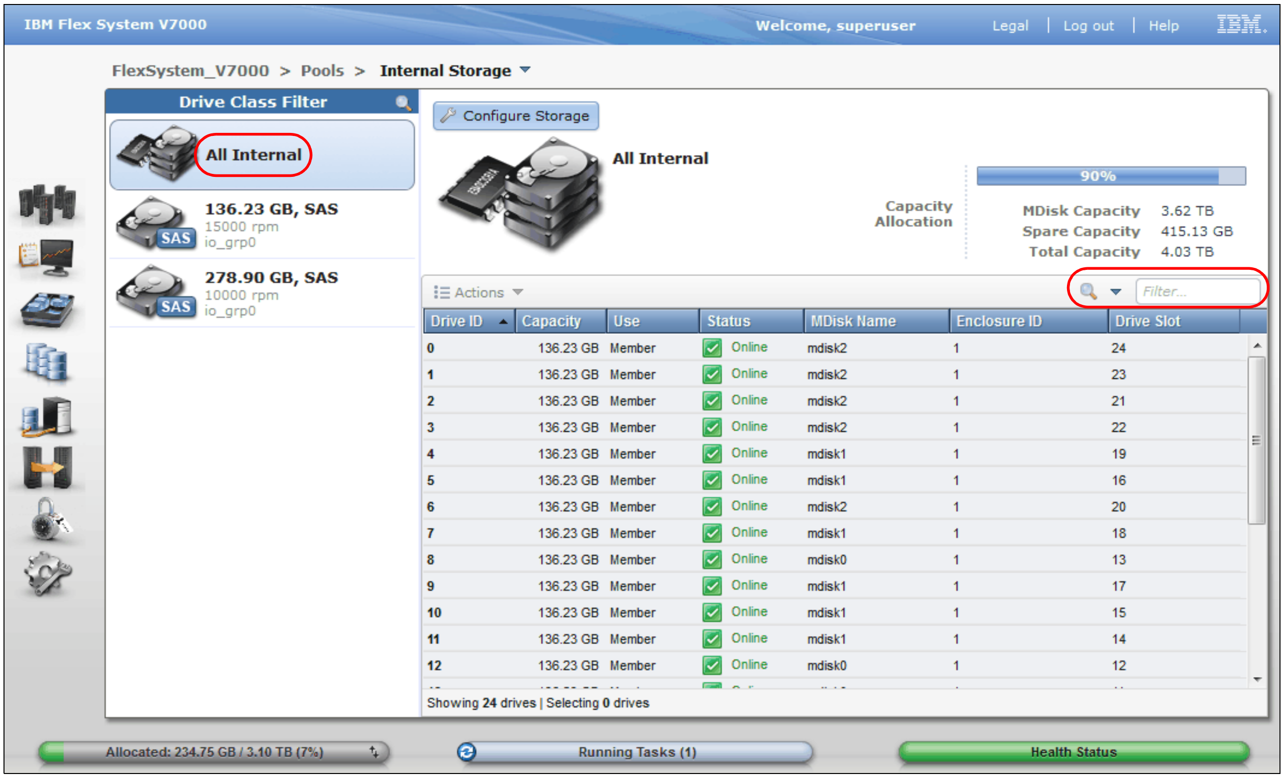


Figure 8-3 Internal Storage window

On the right side of the Internal Storage window, the internal disk drives of the selected type are listed. By default, this additional information is listed along with it:

- Logical drive ID
- Drive’s capacity
- Current type of use (unused, candidate, member, spare, or failed)
- Status (online, offline, and degraded)
- MDisk’s name that the drive is a member of
- Enclosure ID that it is installed in
- Physical Drive Slot ID of the enclosure that it is installed in

More details, for example, the drive's RPM speed or its MDisk member ID, can be shown by right-clicking the blue header bar of the table (Figure 8-4). Also, if you click the table headers, the table content can be ascending or descending sorted by the value represented in that column.

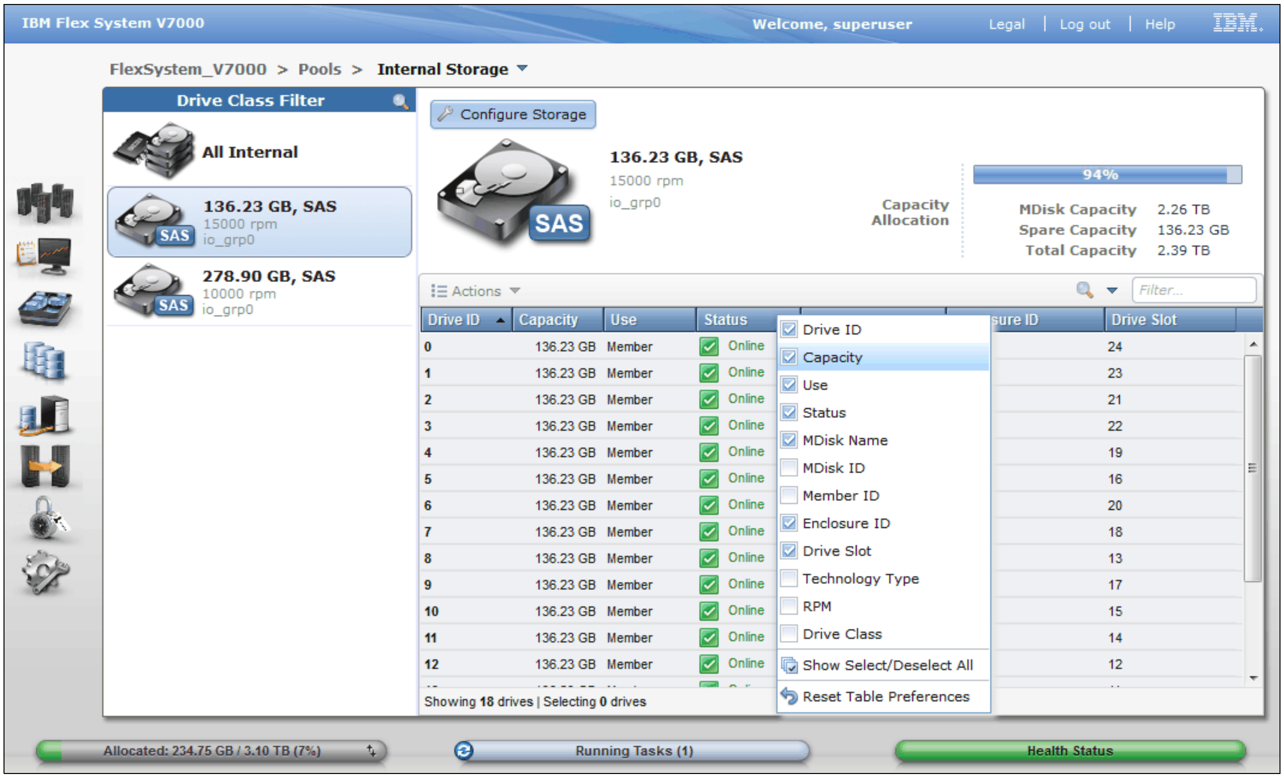


Figure 8-4 Internal Storage window details selection

In addition, you can find the current internal storage capacity allocation indicator at the upper right. The Total Capacity shows the overall capacity of the internal storage installed in this IBM Flex System V7000 Storage Node storage system. The MDisk Capacity shows the internal storage capacity that has been assigned to the MDisks. The Spare Capacity shows the internal storage capacity used for hot spare disks. The percentage bar indicates that 0% capacity has been allocated, because no internal storage has been configured in this example.

8.1.1 Actions on internal drives

There are a few actions that can be taken on internal drives when you select the drive and right-click it, or click the **Actions** drop-down menu (Figure 8-5).

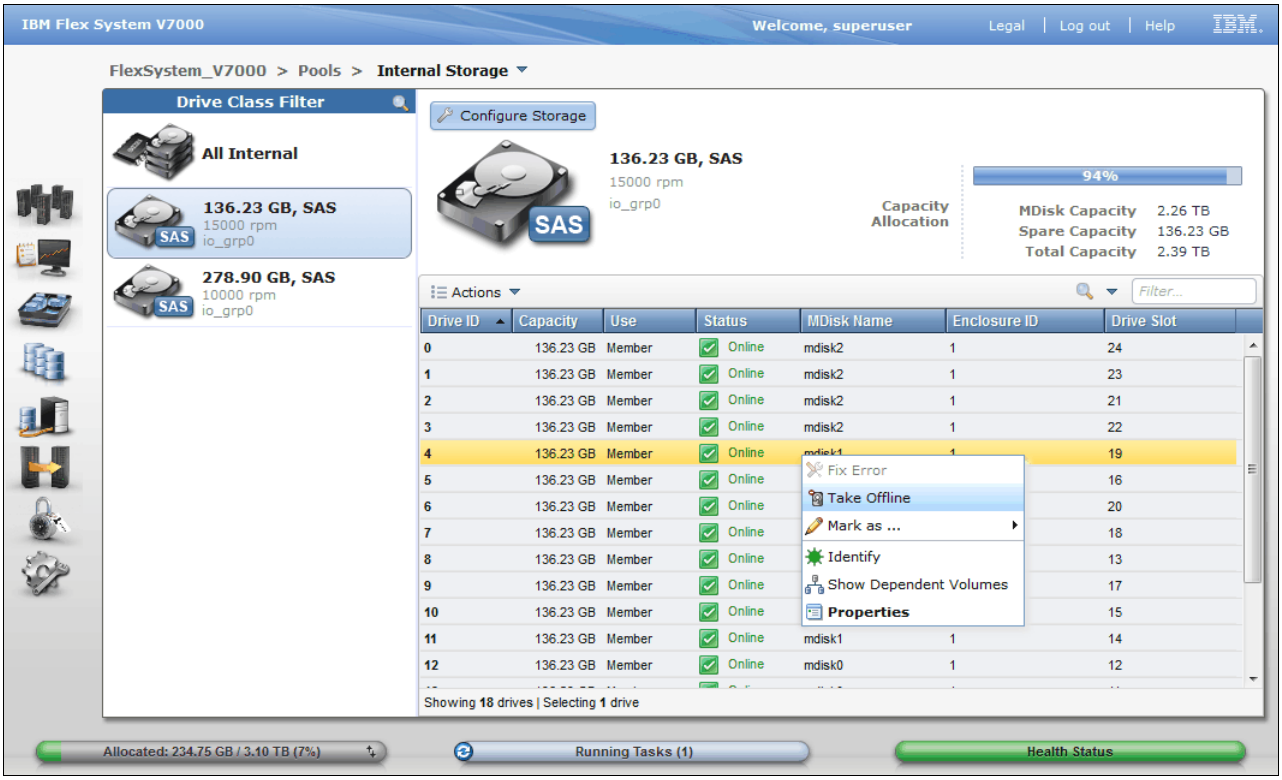


Figure 8-5 Take actions on internal drives

- The Fix Error action starts the Directed Maintenance Procedure (DMP) for a defective drive.
- The internal drives can be taken offline when there are problems on the drives by clicking **Take Offline** in the **Actions** drop-down menu. A confirmation window opens (Figure 8-6). The IBM Flex System V7000 Storage Node storage system prevents the drive from being taken offline if there might be data loss as a result. A drive should be taken offline only if a spare drive is available.

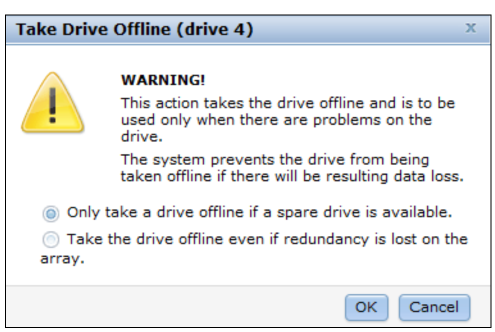


Figure 8-6 Take Drive Offline warning window

Internal drives: Choosing the option to take internal drives offline could lead to potential data loss.

- The internal drives in the IBM Flex System V7000 Storage Node storage system can be assigned to several usage roles, which either can be unused, candidate, spare, or failed. The meanings of these roles are:
  - Unused: The drive is not member of an MDisk (array).
  - Candidate: The drive is available for use in an array.
  - Spare: The drive can be used as a hot spare if required.
  - Failed: The drive was either intentionally taken offline or failed due to an error.

  Click **Mark as...** in the **Actions** drop-down menu, and select the role you want the drive to be assigned to (Figure 8-7).

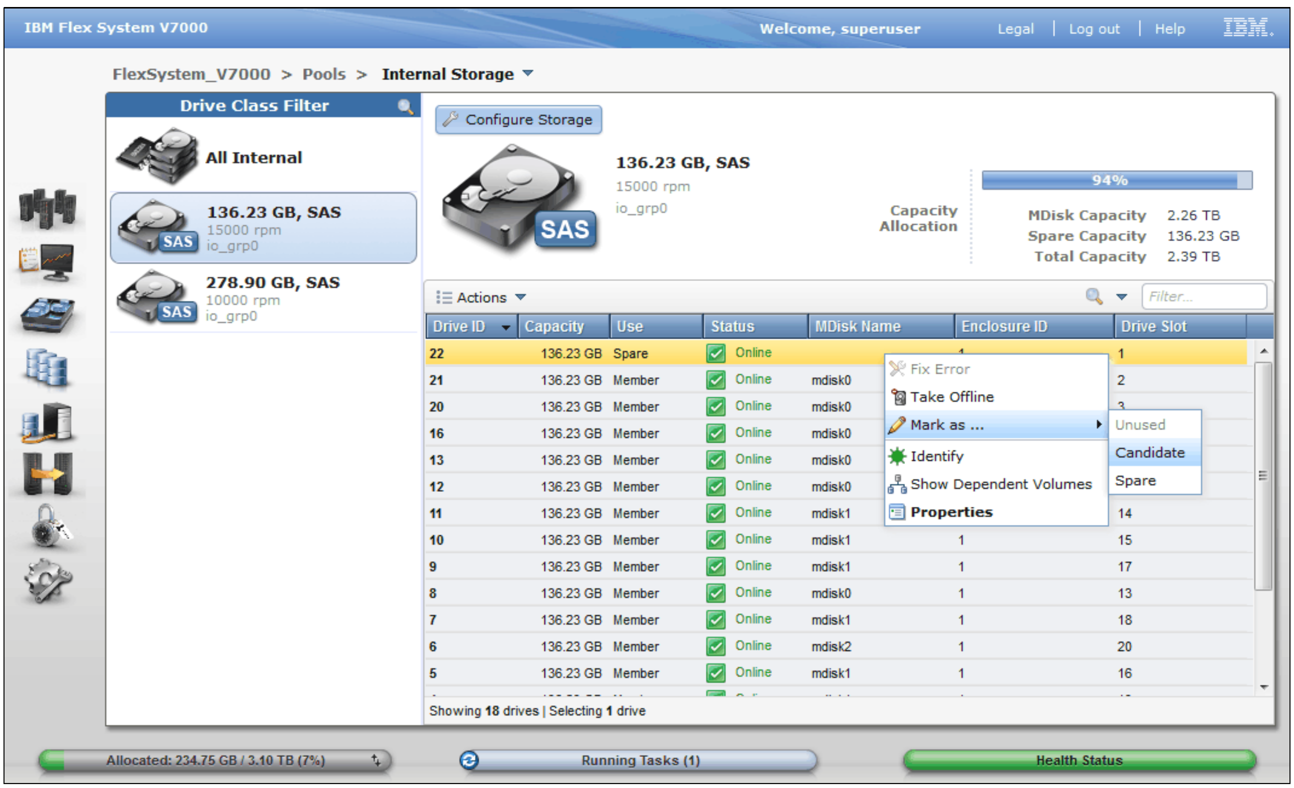


Figure 8-7 Change a drive's usage

- Use the **Identify** action (in Figure 8-7) to turn on the LED light so you can easily identify a drive that needs to be replaced, or that you want to troubleshoot.

Figure 8-8 shows the information that is displayed when you click the **Identify** action.

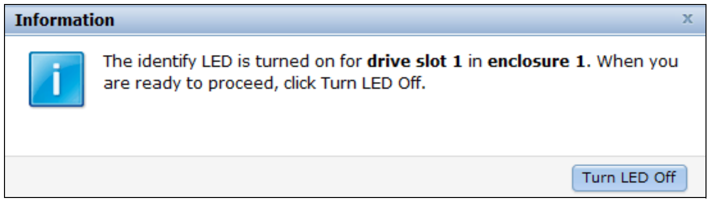


Figure 8-8 Drive identification LED feature

Click **Turn LED Off** when you are done.

- Clicking **Show Dependent Volumes** shows you the volumes that are dependent on the drives. Clicking **Properties** in the **Actions** drop-down menu, or double-clicking the drive, provides more information about the drives (Figure 8-9).

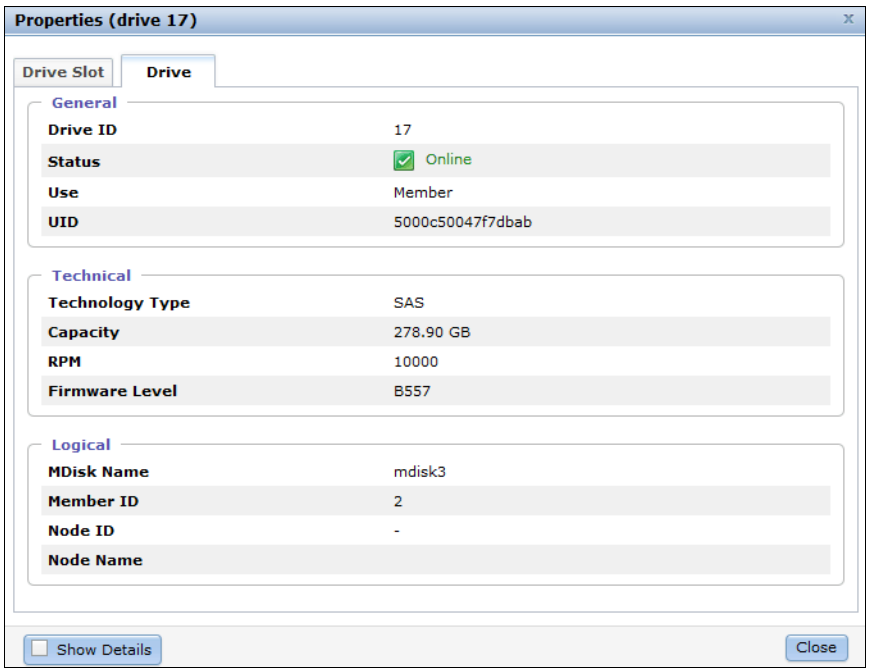


Figure 8-9 Properties tab with default format

- If you select the **Show Details** check box, you can discover more detailed information, including vendor ID, FRU Part Number, and ID (Figure 8-10 on page 303).

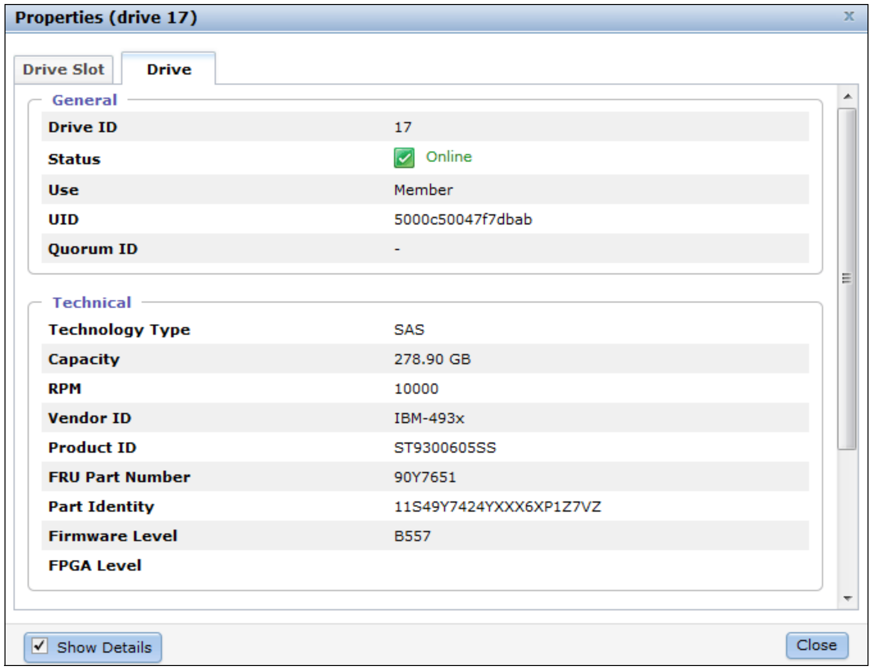


Figure 8-10 Properties tab with details

8.1.2 Configuring internal storage

To configure internal storage for use with hosts, click **Pools** → **Internal Storage** and click the **Configure Storage** button (Figure 8-4 on page 299). A configuration wizard opens and guides you through the process of configuring internal storage (Figure 8-11).

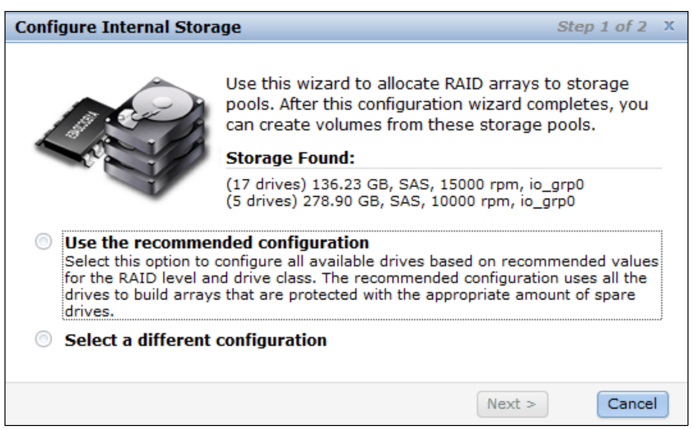


Figure 8-11 Wizard to Configure Internal Storage

The wizard shows all internal drives with a status of candidate available for configuration. If there are internal drives with a status of unused, a window opens, offering to include them in the RAID configuration.

Using the storage configuration wizard simplifies the initial disk drive setup and offers two options:

- Use the recommended configuration
- Select a different configuration

Selecting **Use the recommended configuration** guides you through the wizard described in "Using the recommended configuration" on page 305. Selecting **Select a different configuration** uses the wizard described in "Selecting a different configuration" on page 308.

Before going through the storage configuration wizard, we first describe IBM Flex System V7000 Storage Node RAID configuration presets.

RAID configuration presets

RAID configuration presets are used to configure internal drives based on recommended values for the RAID level and drive class. Each preset has a specific goal for the number of drives per array, the number of spare drives to maintain redundancy, and whether the drives in the array are balanced across enclosure chains, thus protecting the array from enclosure failures. For details about enclosure chains, refer to Chapter 2, "Introduction to IBM Flex System V7000 Storage Node" on page 39.

Table 8-1 describes the presets that are used for solid-state drives (SSDs) for the IBM Flex System V7000 Storage Node storage system.

Table 8-1 SSD RAID presets

Preset	Purpose	RAID level	Drives per array goal	Spare drive goal
SSD RAID 5	Protects against a single drive failure. Data and one stripe of parity are striped across all array members.	5	8	1
SSD RAID 6	Protects against two drive failures. Data and two stripes of parity are striped across all array members.	6	12	1
SSD RAID 10	Protects against at least one drive failure. All data is mirrored on two array members.	10	8	1
SSD RAID 0	Provides no protection against drive failures.	0	8	0
SSD Easy Tier	Mirrors data to protect against drive failure. The mirrored pairs are spread between storage pools to be used for the Easy Tier function.	10	2	1

SSD RAID instances: In all SSD RAID instances, drives in the array are balanced across enclosure chains if possible.

Table 8-2 on page 305 describes the RAID presets that are used for hard disk drives for the IBM Flex System V7000 Storage Node storage system.

Table 8-2 HDD RAID presets

Preset	Purpose	RAID level	Drives per array goal	Spare goal	Chain balance
Basic RAID 5	Protects against a single drive failure. Data and one stripe of parity are striped across all array members.	5	8	1	All drives in the array are from the same chain wherever possible.
Basic RAID 6	Protects against two drive failures. Data and two stripes of parity are striped across all array members.	6	12	1	All drives in the array are from the same chain wherever possible.
Basic RAID 10	Protects against at least one drive failure. All data is mirrored on two array members.	10	8	1	All drives in the array are from the same chain wherever possible.
Balanced RAID 10	Protects against at least one drive or enclosure failure. All data is mirrored on two array members. The mirrors are balanced across the two enclosure chains.	10	8	1	Exactly half of the drives are from each chain.
RAID 0	Provides no protection against drive failures.	0	8	0	All drives in the array are from the same chain wherever possible.

Using the recommended configuration

When creating RAID arrays for the internal drives in IBM Flex System V7000 Storage Node performance and protection against drive failures are important factors to consider. Using the recommended configuration generally builds RAID arrays using RAID 5 which gives good performance with as many drives utilized as possible.

A client or storage administrator may have the need for either higher performance or for more tolerance against failing drives. In such cases you would not use the recommended configuration, but would rather choose a configuration which is optimized for performance as shown in “Selecting a different configuration” on page 308.

As shown in Figure 8-12 on page 306, when you click **Use the recommended configuration**, the wizard offers a recommended storage configuration at the bottom of the window.

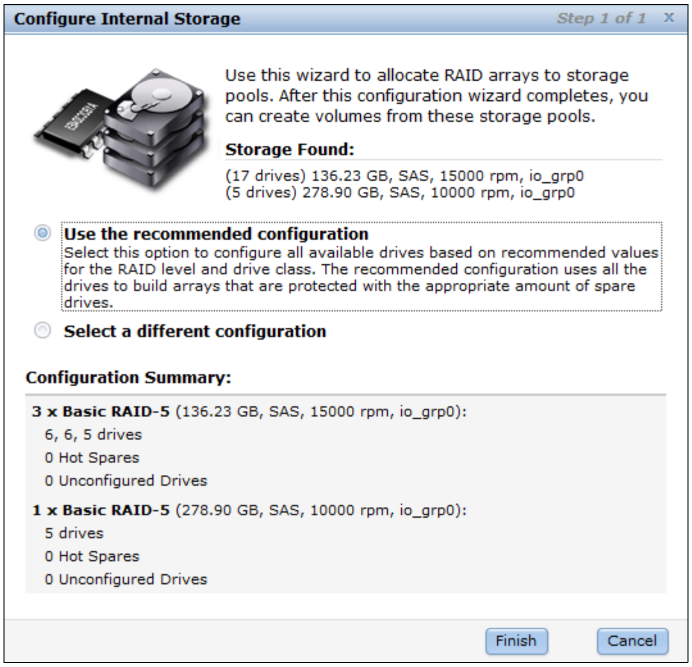


Figure 8-12 Recommended configuration

The recommended RAID presets for different drive classes are:

- SSD Easy Tier preset for solid-state drives
- Basic RAID 5 for SAS drives
- Basic RAID 6 for Nearline SAS drives

Using the recommended configuration, spare drives are also automatically created to meet the spare goals according to the preset chosen; one spare drive is created out of every 24 disk drives of the same drive class on a single chain.

For example, if you have 20 x 450 GB 10 K SAS drives on one chain, one drive in these 20 drives is dedicated as a spare drive. If you have 20 x 450 GB 10 K SAS drives on both chains, which means that there are 10 drives in each chain, then one spare drive on each chain is created. So, if you have 40 x 450 GB 10k SAS drives on both chains, then two spare drives on each chain are created and you have a total of 36 drives that can be array members for the RAID setup.

Spare drives in the IBM Flex System V7000 Storage Node are *global spares*, which means that any spare drive having at least the same capacity as the drive that needs to be replaced can be used in any array. Thus, an SSD array with no SSD spare available would use an HDD spare instead.

In our example above using the recommended configuration the Wizard propose to build three arrays using the Basic RAID 5 for the 146GB drives (discovered as 136,23GB) and one basic RAID 5 for the 300GB drives (discovered as 278,90GB). If the proposed configuration meets your requirements, click **Finish**, and the system automatically creates the array MDisks with a size according to the chosen RAID level. For each of the two different drive types a disk is assigned as hot spare.

Storage pools are also automatically created to contain the MDisks with similar performance characteristics, including the consideration of RAID level, number of member drives, drive class, and so on.

After creating an array, the array MDisk members are being synchronized with each other through a background initialization process. The progress of the initialization process can be monitored by clicking the icon at the left of the **Running Tasks** status bar and then clicking the task displayed (Figure 8-13).

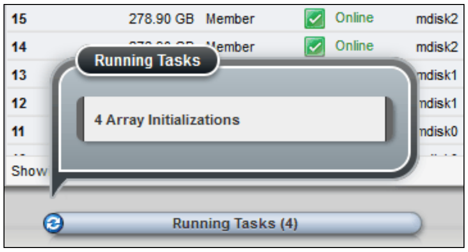


Figure 8-13 Running Tasks status bar

The progress window opens (Figure 8-14). The array is available for I/O during this process. The initialization does not affect the availability due to possible member drive failures.

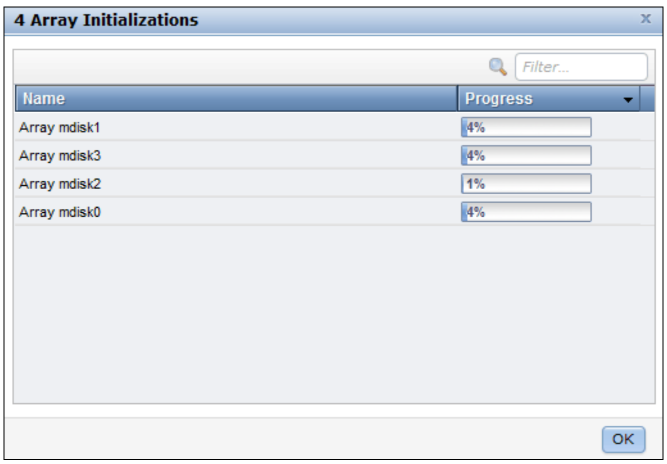


Figure 8-14 Array initialization

The capacity allocation indicator shows that the allocation capacity has reached 90% after the configuration is set (Figure 8-15).

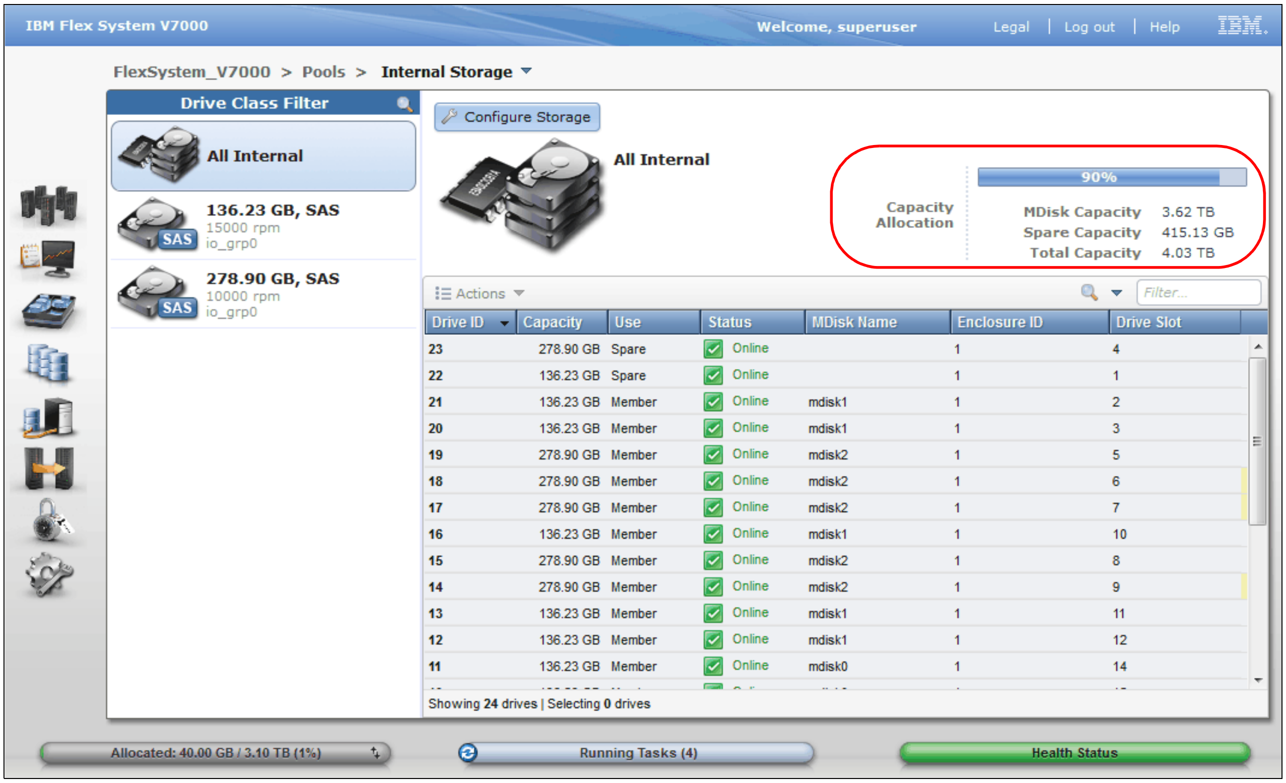


Figure 8-15 Capacity allocation after applying the recommended configuration

If the proposed configuration does not meet your requirements, click **Select a different configuration** on the initial window of the Configure Storage wizard and continue with the more flexible setup shown in "Selecting a different configuration" on page 308.

Selecting a different configuration

The option **Select a different configuration** offers a more flexible way for the configuration of the internal storage as compared to the **Use the recommended configuration** preset in terms of drive selection, RAID level, and storage pool to be used.

Complete the following steps:

1. Click **Select a different configuration** and choose the **Drive Class** you want to configure (Figure 8-16 on page 309).

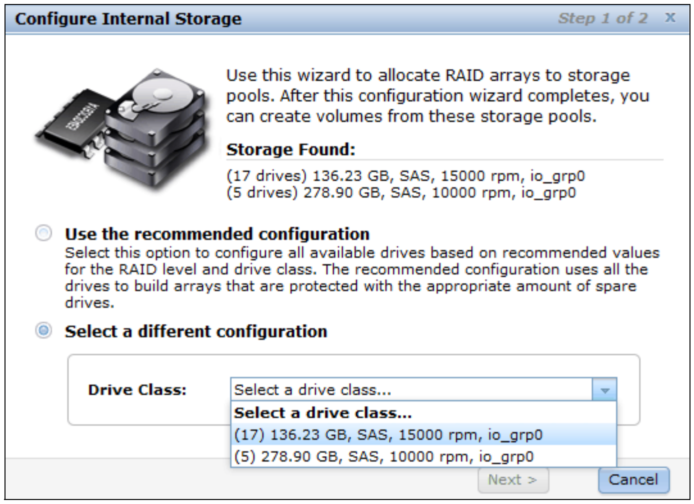


Figure 8-16 Select a drive class

2. Click **Next** and select an appropriate RAID preset (Figure 8-17) and click **Next**.

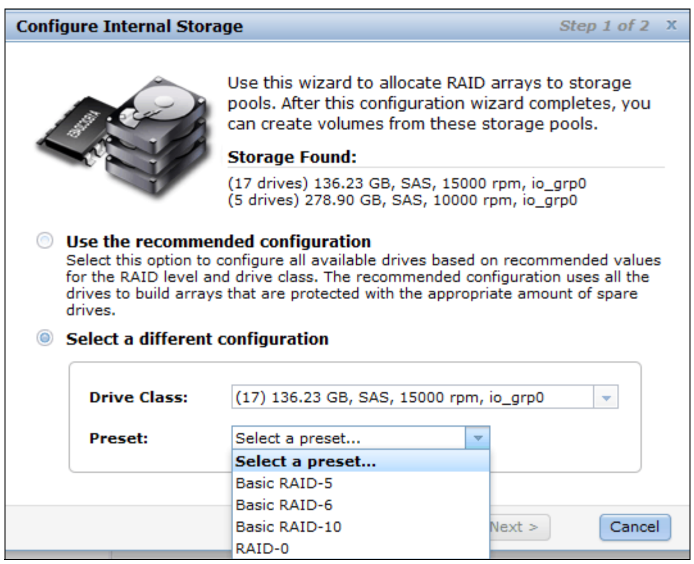


Figure 8-17 Select a RAID preset

3. Select the **Number of drives to provision** in the configuration. By default, all drives with the *Candidate* status are included, and drives with the *Unused* status are made available for use in the array as well. Also, the wizard offers to automatically configure spares to match the spare goal, as explained in "RAID configuration presets" on page 304. Disabling this option enables you to define spare drives manually by marking drives with the status *Candidate* as *Spare*. You then must decide to use one of the two optimization options:

   – Optimize for Performance:

   The goal of this algorithm is to create arrays with identical capacity and performance characteristics to achieve the best possible performance for the volumes to be created. In a performance optimized setup, the IBM Flex System V7000 Storage Node provisions eight physical disk drives in a single array MDisk, except for the following situations:

   - RAID 6 uses 12 disk drives.

- SSD Easy Tier uses two disk drives.

Hence, creating an *Optimized for Performance* configuration is only possible if sufficient number of drives available to match your needs.

As a consequence, all arrays with similar physical disks feature the same performance characteristics. Due to the defined presets, this setup might leave drives unused. The remaining unconfigured drives can be used then in another array.

Figure 8-18 shows an example of the performance optimized setup.

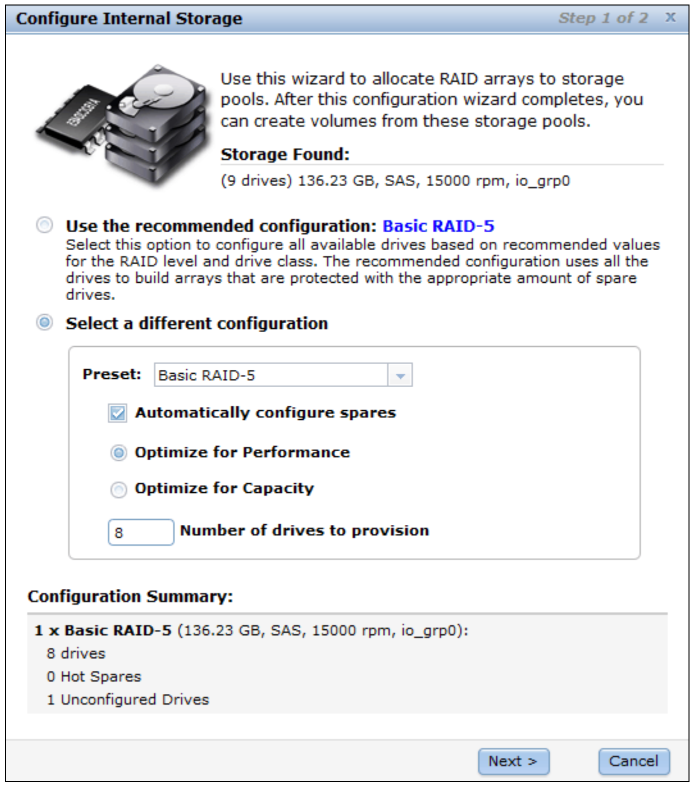


Figure 8-18 Example of a performance optimized setup

Choose the storage pool you want to assign the capacity to (Figure 8-19 on page 311). You can either select an existing storage pool that does not contain MDisks, or a pool containing MDisks with similar performance characteristics, which is listed automatically.

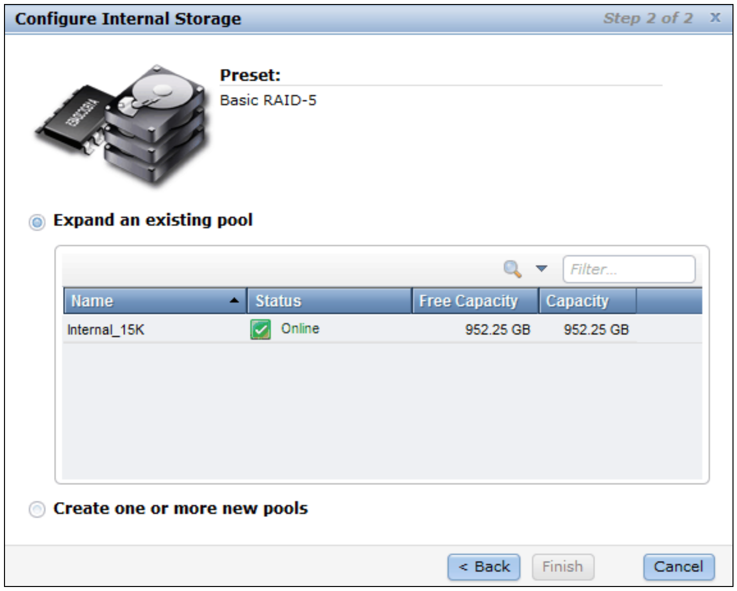


Figure 8-19 Assign capacity to an existing storage pool

Alternatively, create a new storage pool and type in a name for the new storage pool (Figure 8-20).

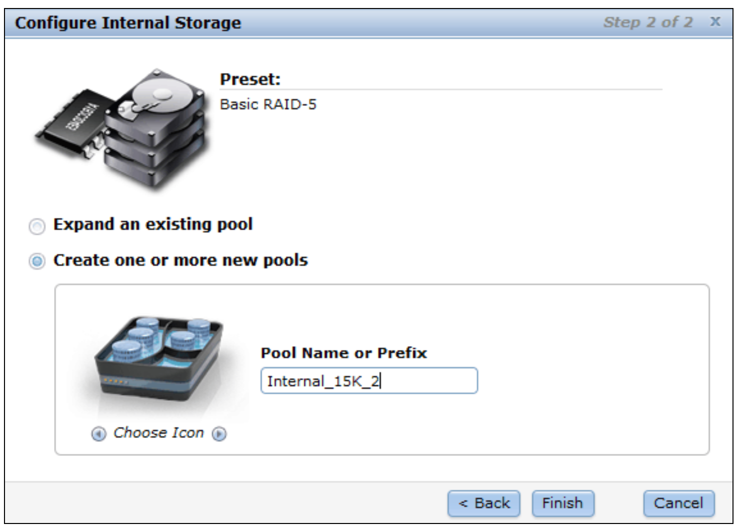


Figure 8-20 Create storage pool

Click **Finish** to finalize the wizard. After the wizard completes, the configuration changes are reflected at the upper right of the Internal Storage window (Figure 8-21 on page 312).

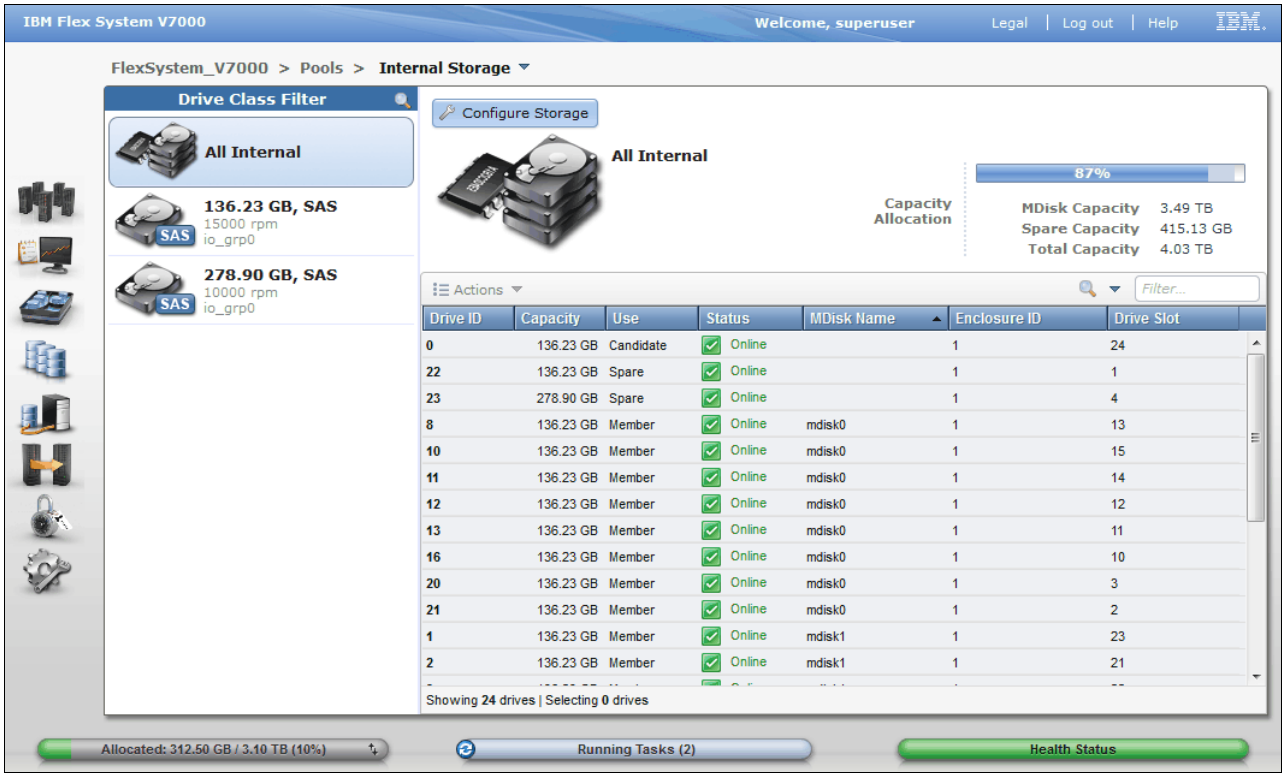


Figure 8-21 Configuration complete with performance optimized setup

With the performance optimized setup the IBM Flex System V7000 Storage Node creates MDisks, each containing eight drives. Additional drives remain unconfigured, as they did not meet the goal of eight drives for a Basic RAID 5 array to create another *performance optimized* MDisk. The progress of the array initialization process is displayed in the Running Tasks status indicator, and the capacity allocation indicator has been updated as well.

– Optimize for Capacity:

The goal of this algorithm is to create a setup with the maximum usable capacity, depending on the selected RAID level.

While creating arrays, the IBM Flex System V7000 Storage Node system attempts to fulfill the *width goal* (Table 8-2 on page 305) for each array before creating another one. This setup is an example with 18 *unused* drives, **Select a different configuration** was chosen with the following options:

- Preset: Basic RAID-5.
- Automatically configure spares was checked.
- Optimize for capacity.

The results are as follows:

- One hot spare drive is defined.
- Three basic RAID 5 arrays are created. Two of the arrays contain six drives and one array contains the remaining five drives.
- No unused drives are left.

The *width goals* for the array can not be accomplished as too little drives are available. Instead the system provisions the available drives in even sized MDisks. In our case we get two MDisks with six disks and one with only five.

In the Capacity Optimized setup, all available Candidate disk drives are used, no "unconfigured drives" remain, unlike in the Performance Optimized setup. Figure 8-22 shows an example of the capacity optimized setup.

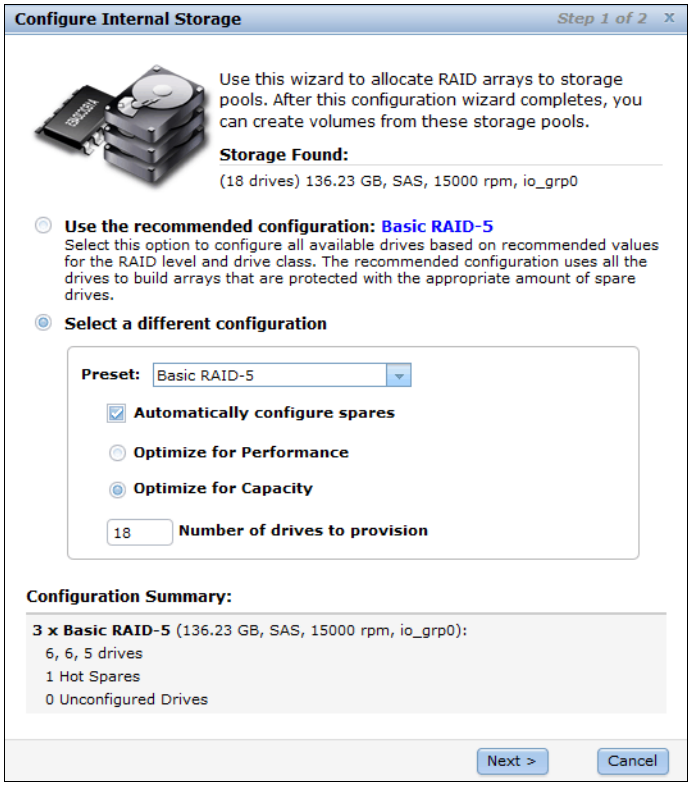


Figure 8-22 Capacity optimized setup

Provisioned drives: You can choose the number of drives to be provisioned. The wizard proposes that you create arrays from all available candidate drives. If you want to configure only a subset of drives at this moment in time, you can change the number in the Drives to provision field to the wanted value. If you want to manually select the drives to provision, assign only the wanted drives Candidate status. Alternatively, you can use the CLI for full control over the drive configuration, which is beyond the intended scope of this book.

8.2 Working with MDisks

After the configuration is completed for the internal storage, you can find the MDisks created on the internal arrays in the MDisks by Pools window.

You can access the MDisks window by clicking **Home** → **Overview** and clicking the **MDisks** function icon. In the extended help information window, click **Pools**.

Figure 8-23 on page 314 shows how to access the **Pools** window from the Getting Started window.

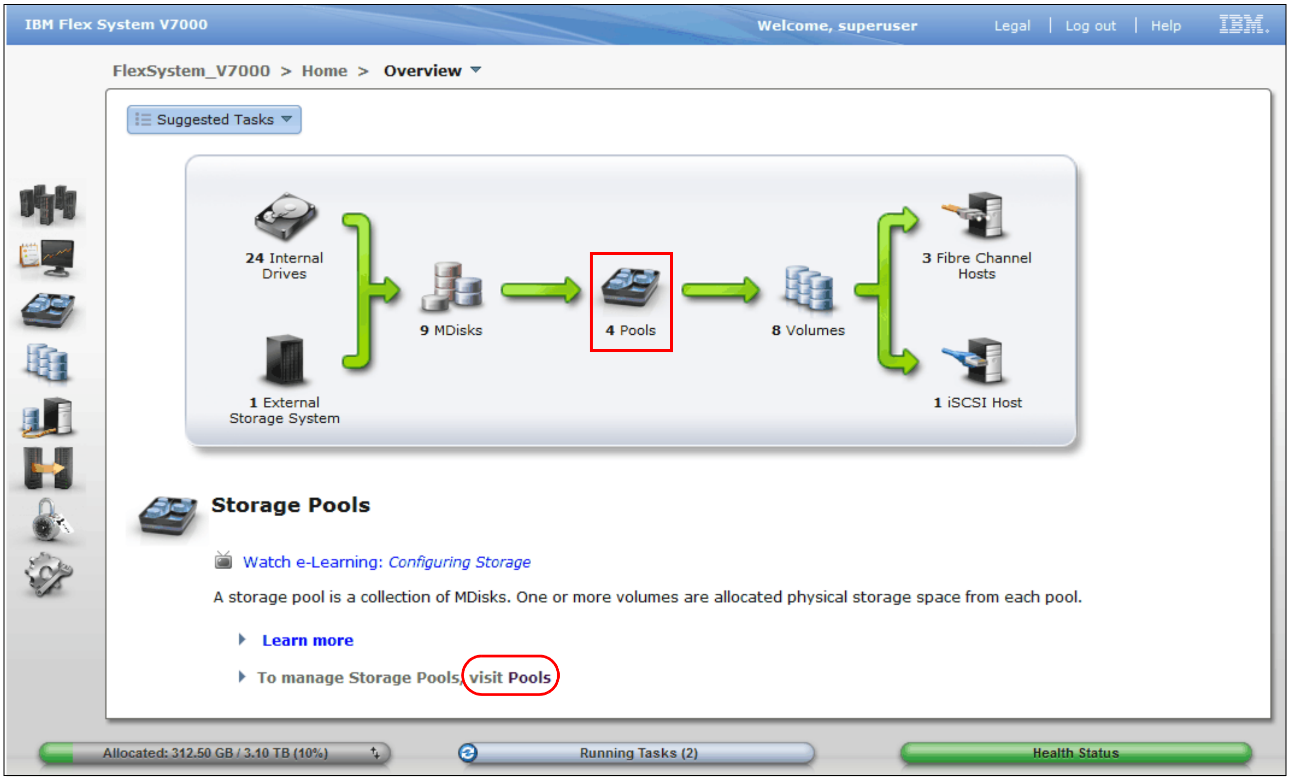


Figure 8-23 Access the Pools management window from the Overview window

Another way to access the MDisks window is by using the Physical Storage function icon on the left (Figure 8-24).

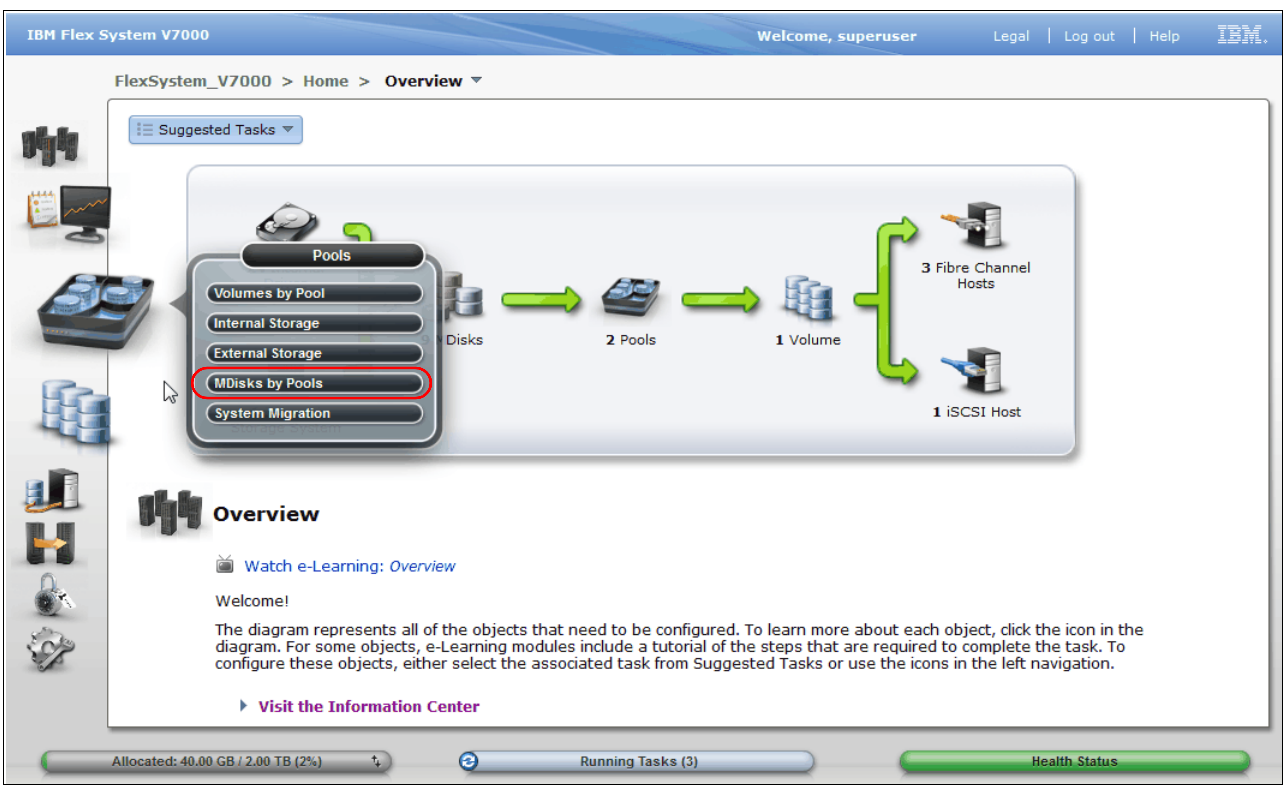


Figure 8-24 Access the MDisks by Pools window using the Physical Storage function icon

The MDisks by Pools window (Figure 8-25) lets you manage all MDisks made of both internal and external storage.

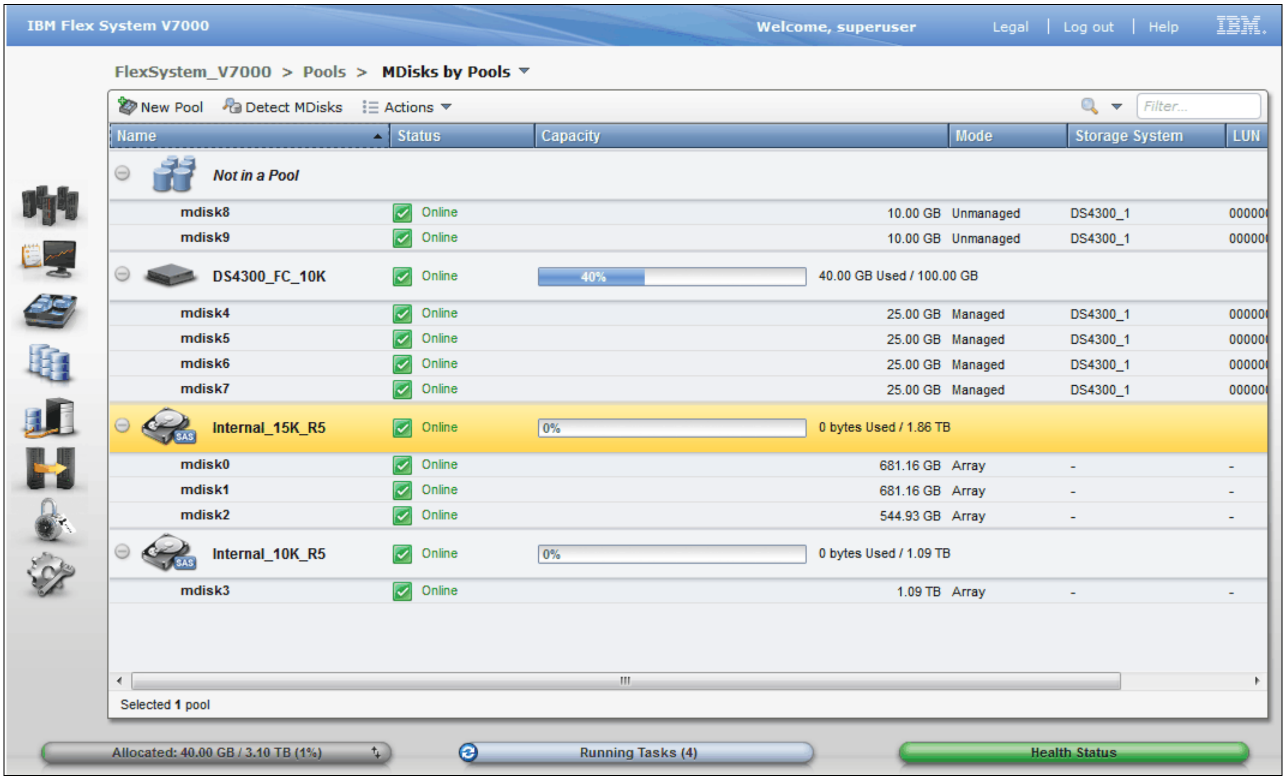


Figure 8-25 MDisks by Pools window

The window provides the following information:

- MDisk name
- Numerical ID
- Status
- Capacity
- Mode
- Name of the storage pool it belongs to
- Name of the backing storage system for MDisk on external storage
- MDisk's LUN ID from external storage systems
- Assigned storage tier

Not all the above values may show by default, but can be enabled by right-clicking the blue bar above the MDisk groups.

You can find more information about attaching external storage to an IBM Flex System V7000 Storage Node storage system in Chapter 11, "SAN Connections and Configuration" on page 431.

In the IBM Flex System V7000 Storage Node there are four modes for MDisks:

- Array: Array mode MDisks are constructed from internal drives using the RAID functionality. Array MDisks are always associated with storage pools (also known as MDisk Groups).
- Unmanaged: The MDisk is not a member of any storage pools, which means it is not being used by the IBM Flex System V7000 Storage Node storage system. LUNs presented by

external storage systems to IBM Flex System V7000 Storage Node are discovered as unmanaged MDisks.

- Managed: The MDisk is assigned to a storage pool and provides extents to be used by volumes.
- Image: The MDisk is assigned directly to a volume with a one-to-one mapping of extents between the MDisk and the volume.

8.2.1 Adding MDisks to storage pools

By adding *unmanaged MDisks* to a pool, their status changes to *managed MDisks*. Managed MDisks can belong to only one pool. Unmanaged MDisks can either be added to a new pool or to an existing pool to expand its capacity. Pools are commonly used to group MDisks from the same storage subsystem and with the same class of speed.

A new pool can be created in the MDisks by Pools window by clicking the **New Pool** icon. Assign a name to the pool and choose an icon, if wanted (Figure 8-26).

Figure 8-26 Create storage pool - step 1

The *Create Pool* window (Figure 8-27) lets you include unmanaged MDisks in the new pool. Several filter options at the top of the window allow you to limit the selection, for example, by storage subsystem, capacity, and so on. Several MDisks can be selected by pressing the Ctrl key while clicking the MDisks listed with the mouse. Also, the Detect MDisks icon initiates a SAN discovery for finding recently attached external storage systems.

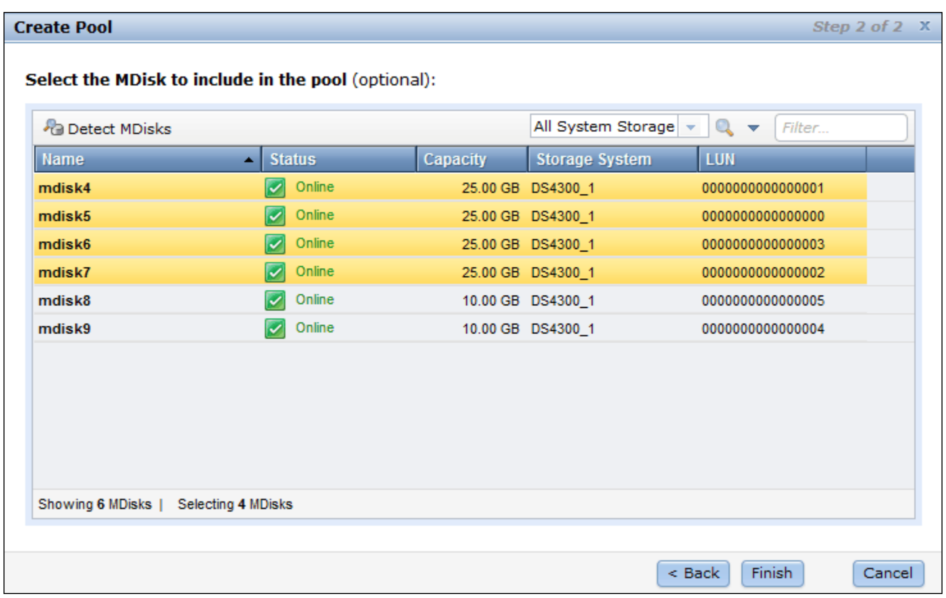


Figure 8-27 Create storage pool - step 2

To add an unmanaged MDisks to an existing pool select the MDisk, click **Actions** → **Add to Pool** (Figure 8-28).

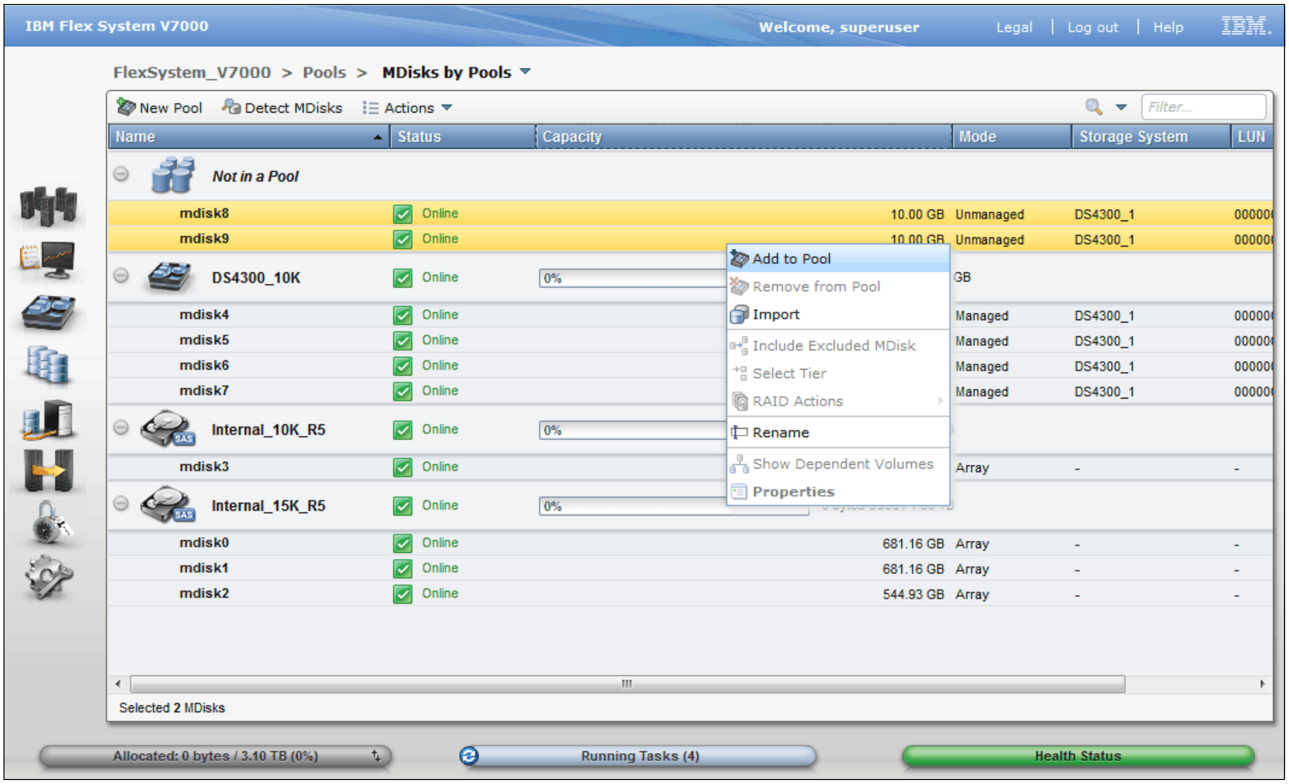


Figure 8-28 Add an unmanaged MDisk to a storage pool

> **Existing data:** If there is existing data on the unmanaged MDisks that you need to preserve, do *not* select **Add to Pool** on this LUN, as this action deletes the data. Use **Import** instead, which is described in 8.2.2, "Importing MDisks" on page 320.

Choose the storage pool to which you want to add the MDisk and click **Add to Pool** (Figure 8-29).

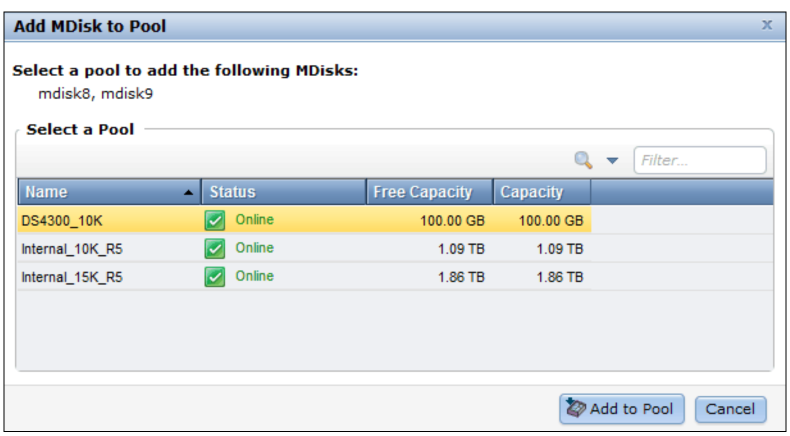


Figure 8-29 Add MDisk to Pool

After the IBM Flex System V7000 Storage Node system completes this action, the MDisk shows up in the pool that it was added to (Figure 8-30).

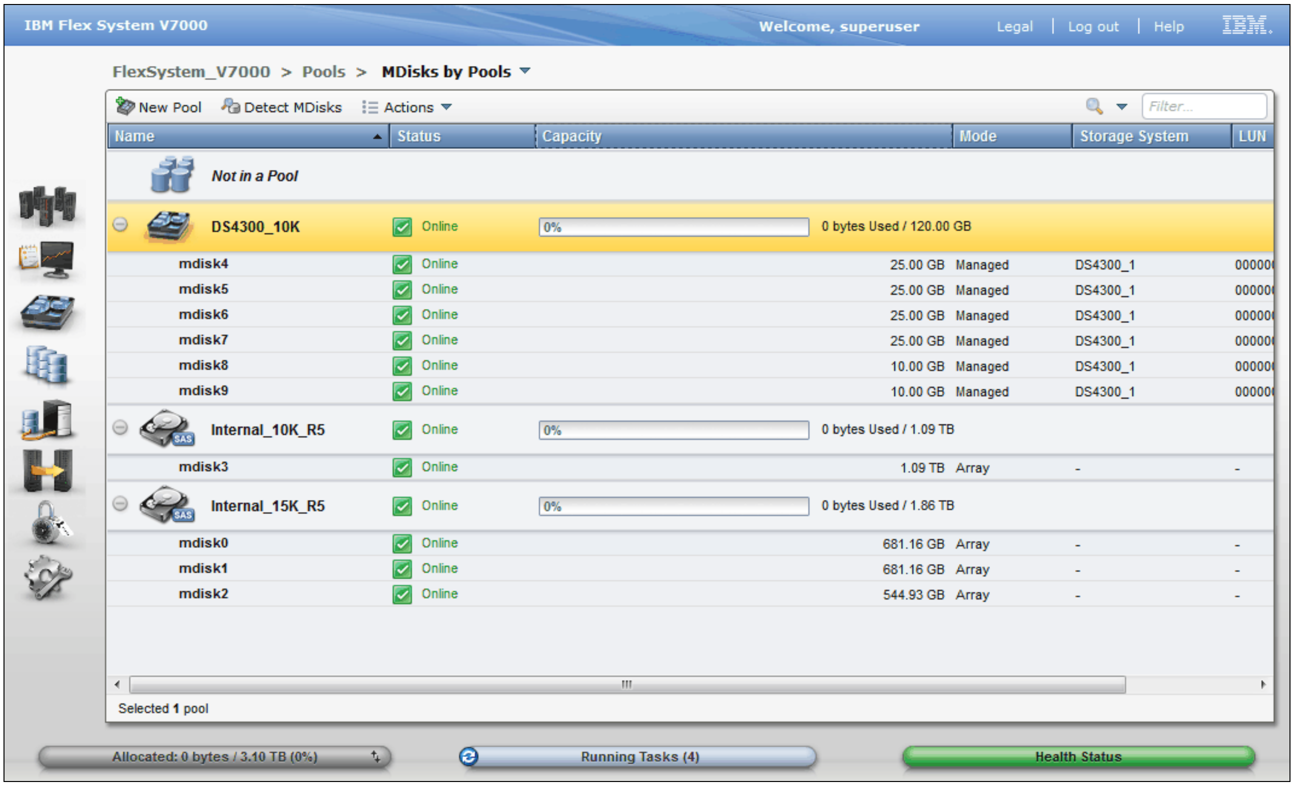


Figure 8-30 MDisk added to pool

In some cases, you may want to remove MDisks from storage pools to reorganize your storage allocation. You can remove MDisks from storage pools by selecting the MDisks and clicking **Remove from Pool** from the **Actions** drop-down menu (Figure 8-31 on page 319).

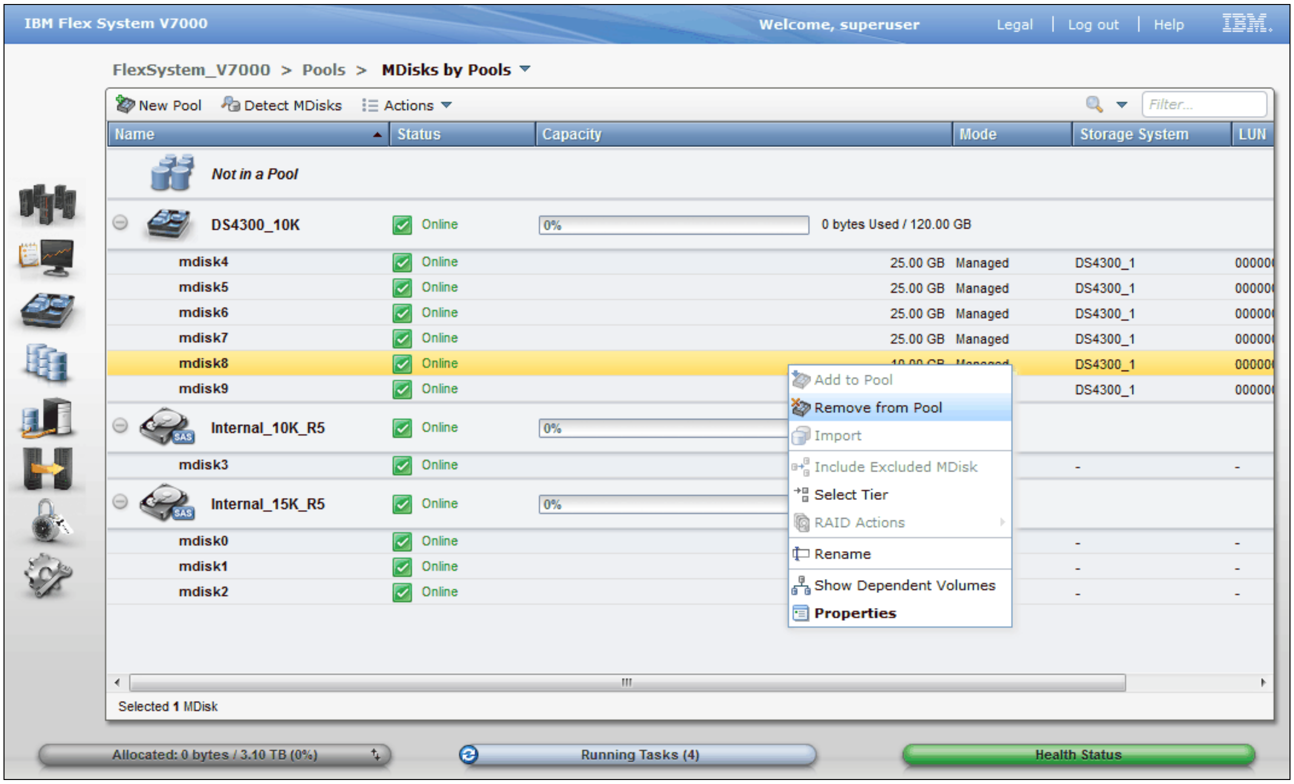


Figure 8-31 Remove an MDisk from the storage pool

You need to confirm the number of MDisks you need to remove (Figure 8-32). If you have data on the MDisks, and you still need to remove the MDisks from the pool, select the **Remove the MDisk from the storage pool even if it has data on it. The system migrates the data to other MDisks in the pool** check box.

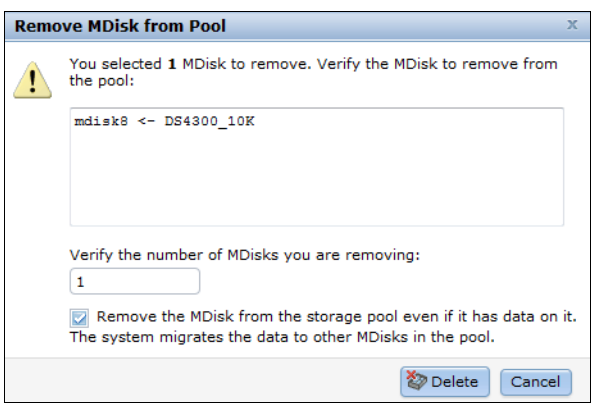


Figure 8-32 Confirm the removal of MDisk from the pool

Available capacity: Make sure that you have enough available capacity left in the storage pool for the data on the MDisks to be removed.

After you click **Delete**, data migration from the MDisks to be removed starts. You can find the migration progress in the Running Tasks status indicator (Figure 8-33).

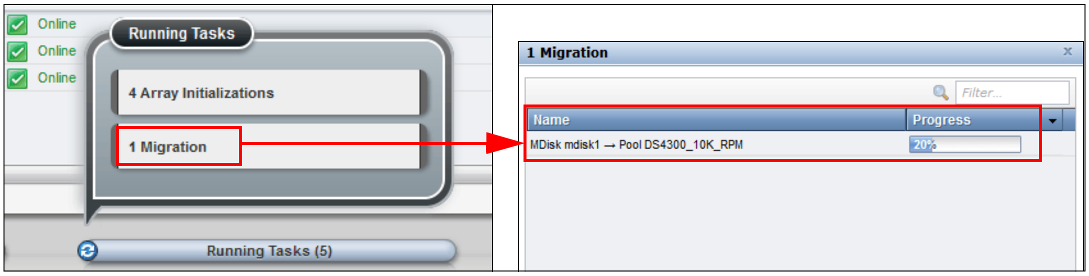


Figure 8-33 Data migration progress when removing MDisks from the pool

8.2.2 Importing MDisks

LUNs hosted on external storage systems can be migrated into IBM Flex System V7000 Storage Node storage. Hosts, which used to be directly attached to those external storage systems, can continue to use their storage that is now presented by the IBM Flex System V7000 Storage Node instead.

To achieve this configuration, the existing external LUNs must be imported as an *image-mode volume* using the Import option. This action is possible for unmanaged MDisks only. Those disks must not have been added to a pool, as described in 8.2.1, “Adding MDisks to storage pools” on page 316.

If the Import option is used and no existing storage pool is chosen, a temporary *migration pool* is created to hold the new image-mode volume. This image-mode volume has a direct block-for-block translation from the imported MDisk to the volume presented from now on by the IBM Flex System V7000 Storage Node and existing data is being preserved.

In Figure 8-34, we show an example of how to import an unmanaged MDisk. Select the unmanaged MDisk and click **Import** from the **Actions** drop-down menu or right-click the MDisk and select **Import**.

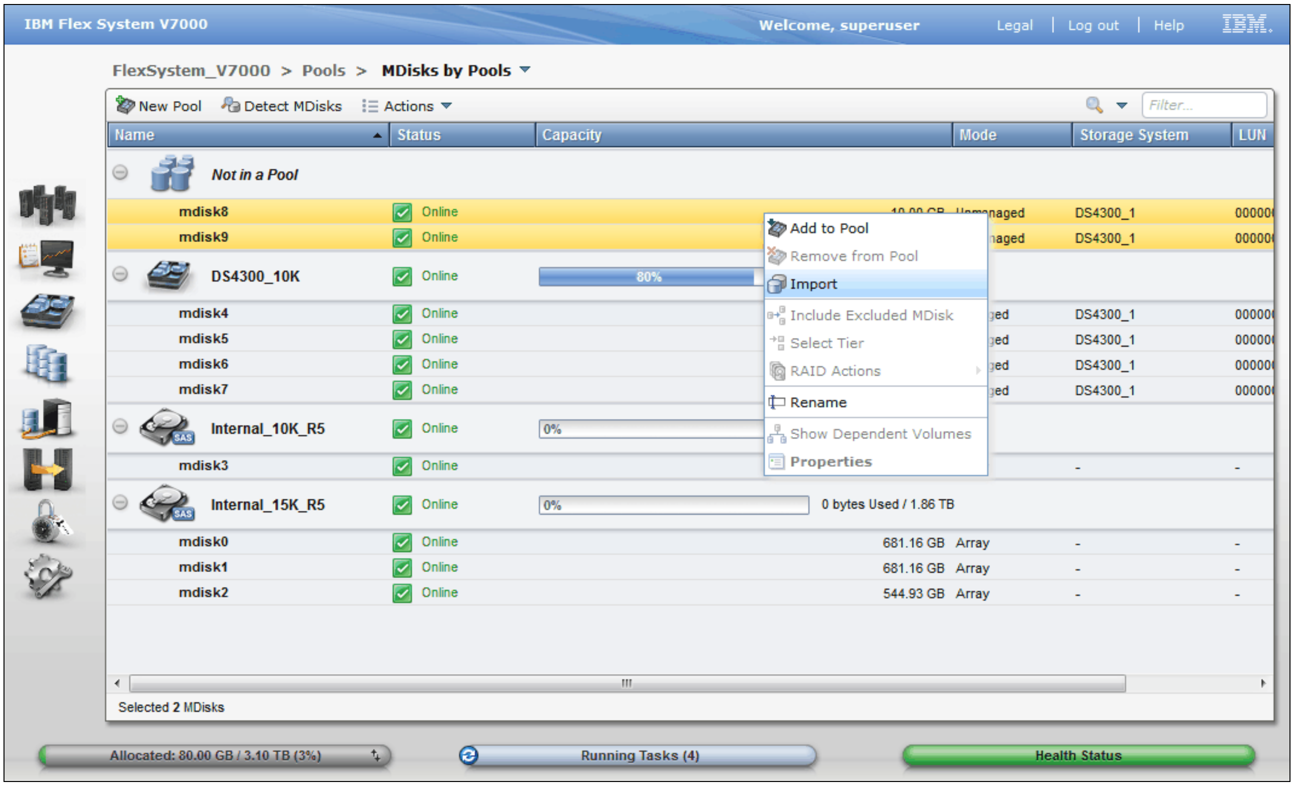


Figure 8-34 Import MDisk

As shown in Figure 8-35, the import wizard is activated and guides you through the import process.

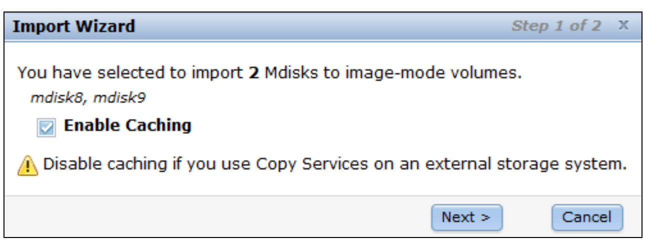


Figure 8-35 Import Wizard - step 1 of 2

In Step 1 of the Import Wizard, you must decide whether the volume is imported as a generic volume (the default) or as thin-provisioned. Caching for the volume can be disabled; it is enabled by default.

- Select **Use thin-provisioning** only if the volume to be imported originally was a thin-provisioned volume that was exported to an image-mode volume.
- Clear **Enable Caching** if you use copy services on the external storage system hosting the LUN. It is a best practice to use the copy services of IBM Flex System V7000 Storage Node for volumes virtualized. You can find more information about virtualizing external storage in Chapter 7, “Migration wizard” on page 265.

If you want to migrate the data from the external volume (that is being imported) to existing IBM Flex System V7000 Storage Node MDisks, choose an existing destination storage pool. Only pools with sufficient free capacity are listed in Step 2 (Figure 8-36). The actual data migration begins after the MDisk was imported successfully.

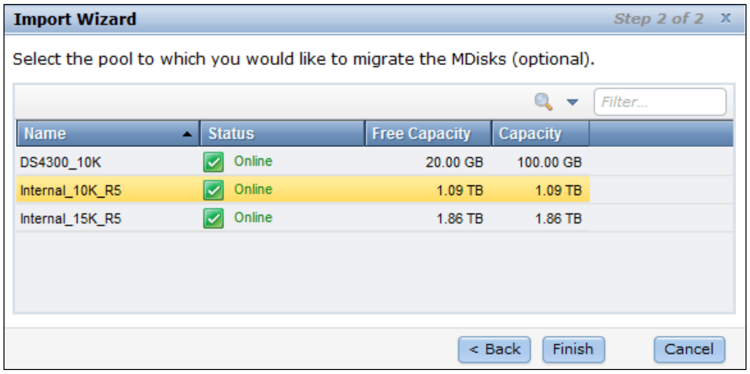


Figure 8-36 Import Wizard - step 2 of 2

You can check the migration progress in the Running Tasks status indicator as also shown in Figure 8-33 on page 320 or by clicking **Pools** → **System Migration**.

After the migration has completed, you can find the volume in the chosen destination pool (Figure 8-37).

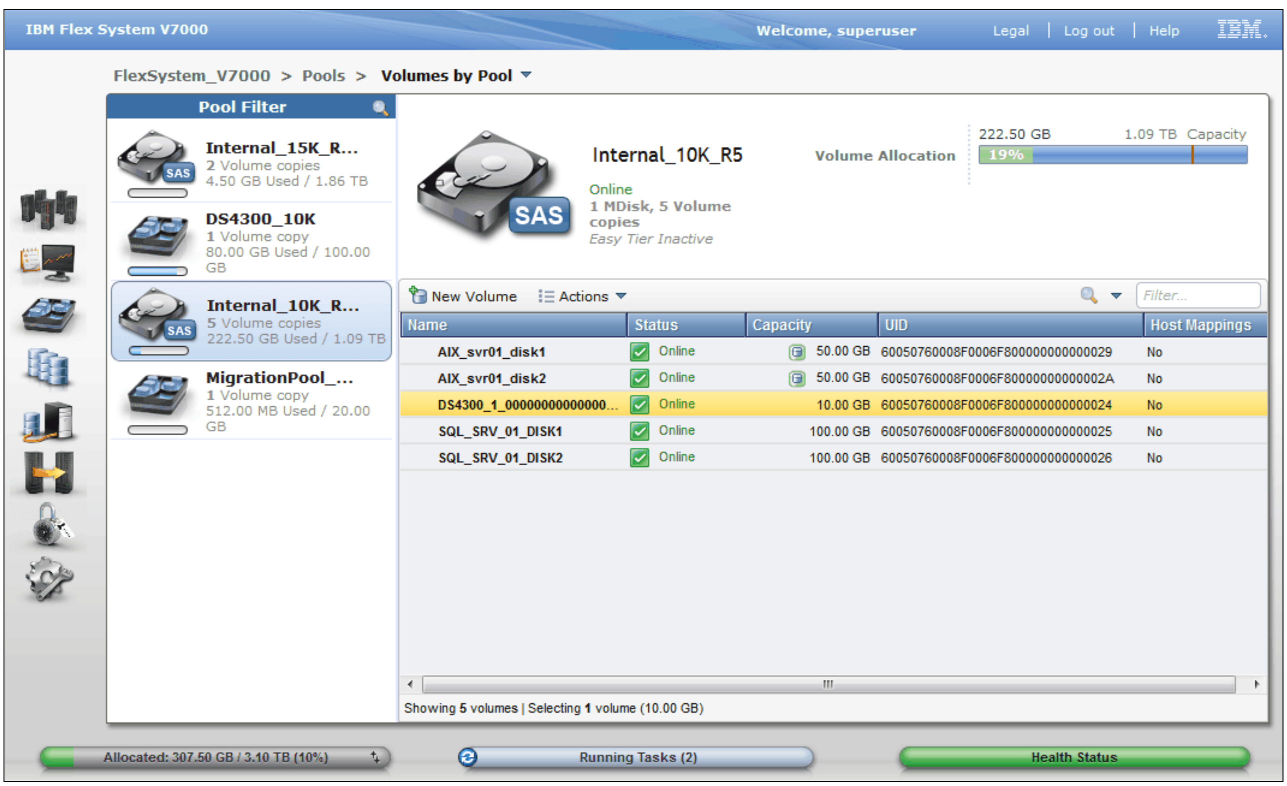


Figure 8-37 Volume migrated to destination pool

The name of the Volume is determined by the IBM Flex System V7000 Storage Node MDisk Import process, and you may want to rename the Volume to a more adequate name that indicate the purpose of the Volume. The volume can be renamed by right-clicking the volume and click **Rename**.

All data has been migrated off the source MDisk to MDisks in the destination storage pool. The source MDisk has changed its status to managed and is associated with an automatically created migration pool. It can be used as a regular MDisk to host volumes (Figure 8-38).

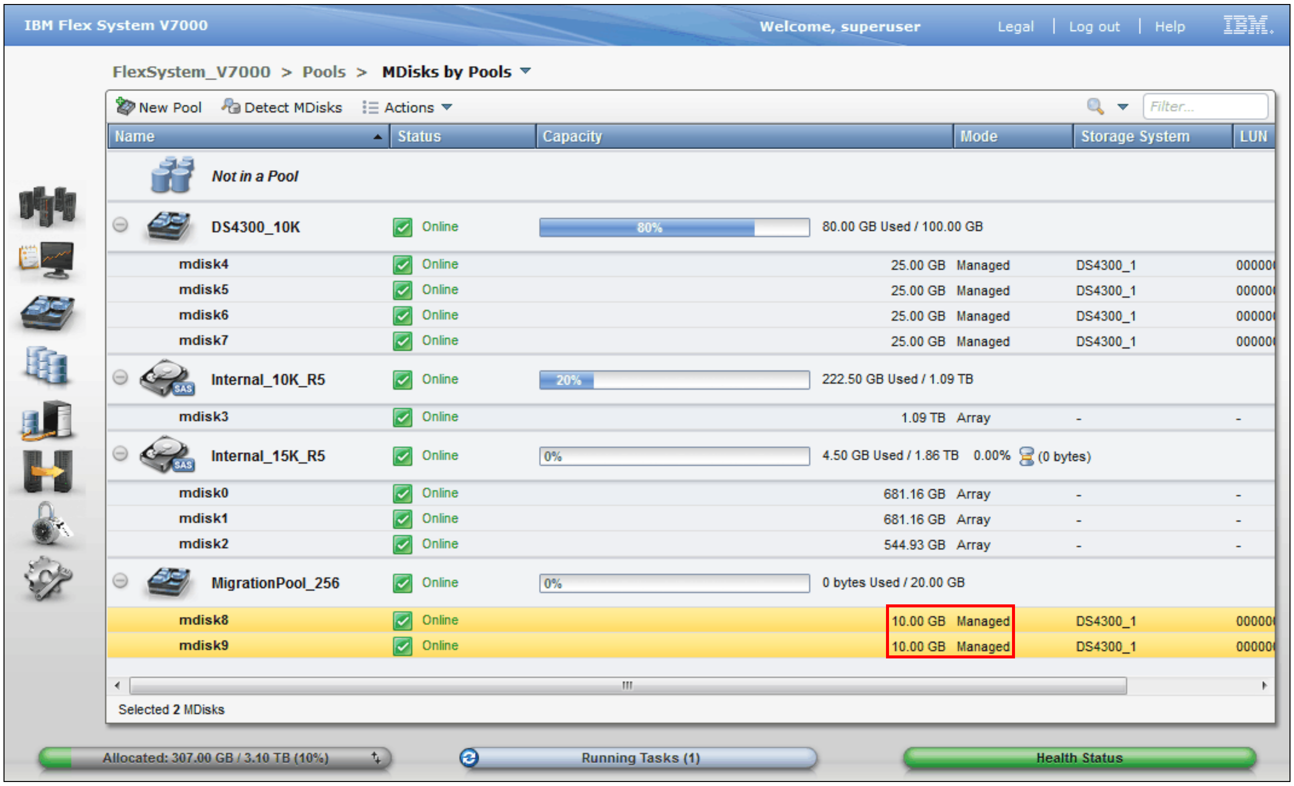


Figure 8-38 MDisk mode changed to Managed after a destination pool was selected during Import

If you do not select a destination pool during Step 2 of the Import Wizard and click **Finish**, a warning window opens (Figure 8-39). The imported MDisk remains in its temporary storage pool as an image mode volume.

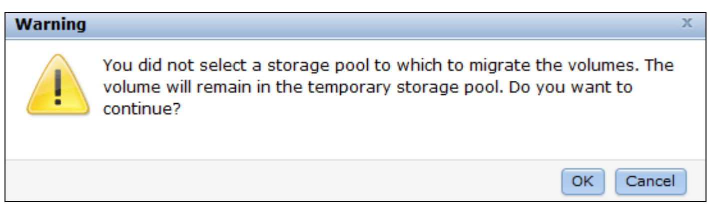


Figure 8-39 Warning window that opens if no destination storage pool was chosen

After the import process has completed, the MDisk shows up as image mode volume (Figure 8-40 on page 324).

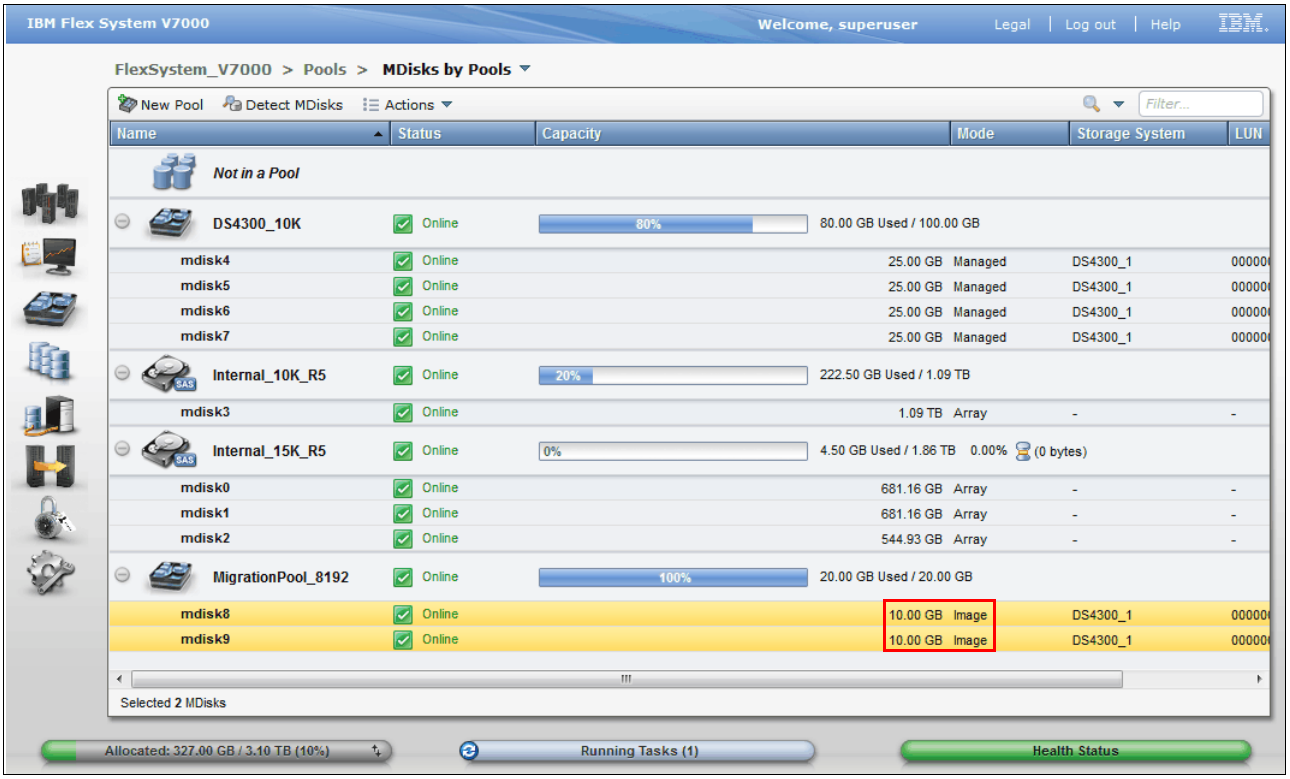


Figure 8-40 MDisk mode changed to Image when no destination pool was selected during Import

If needed, the image mode volume can be migrated manually into a different pool by selecting either **Migration to Another Pool** or **Volume Copy Actions**.

Alternatively, the migration into another pool can be done by clicking **Pools** → **System Migration**. Migration is described in greater detail in Chapter 7, “Migration wizard” on page 265.

Any imported MDisk that has not been migrated into a pool yet is listed under **Pools** → **System Migration** (Figure 8-41).

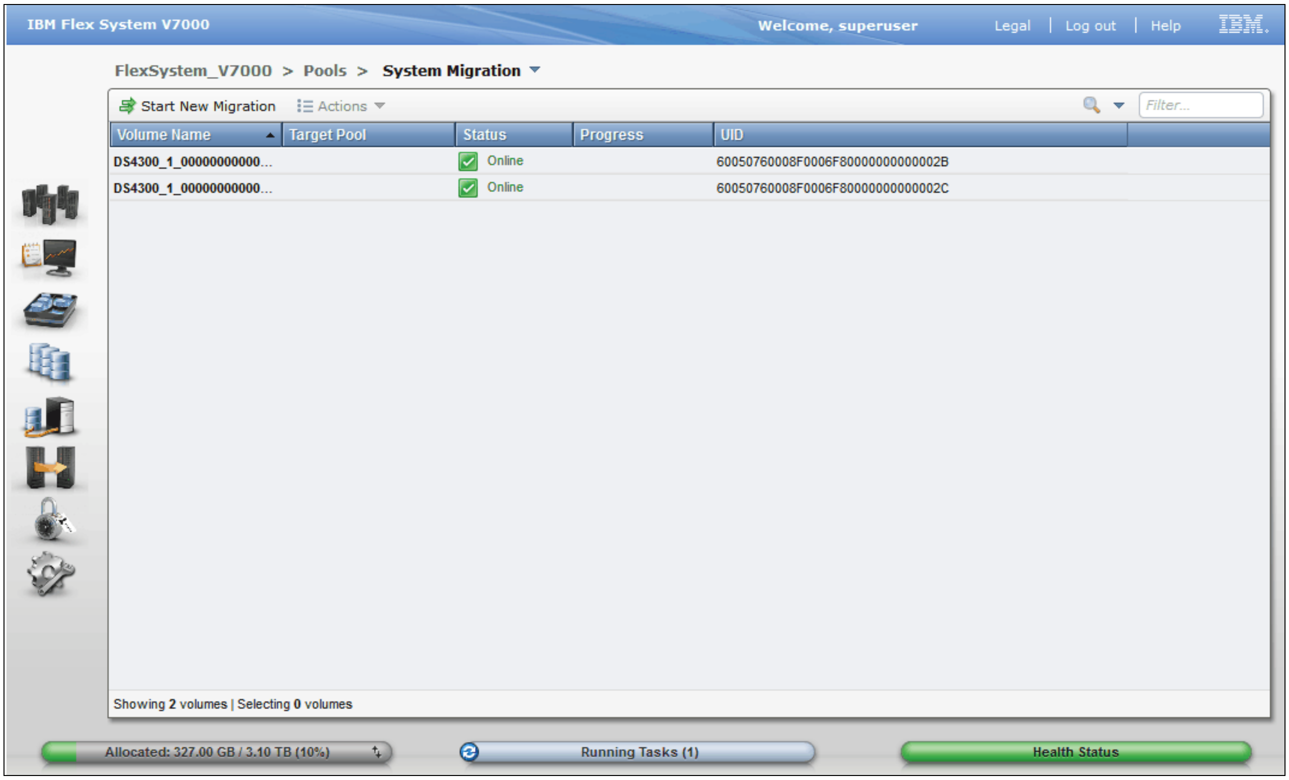


Figure 8-41 Imported MDisk in the System Migration window

This feature is normally used as a method to migrate data from existing external LUNs into storage pools, either located internally or externally, on the IBM Flex System V7000 Storage Node. You should not use image mode volumes as a *long-term solution* for reasons of performance and reliability.

To migrate an image mode volume into a regular storage pool, select the volume to be migrated and click **Actions** → **Start Migration**.

Figure 8-42 on page 326 shows that the migration of the Image Mode disk is started.

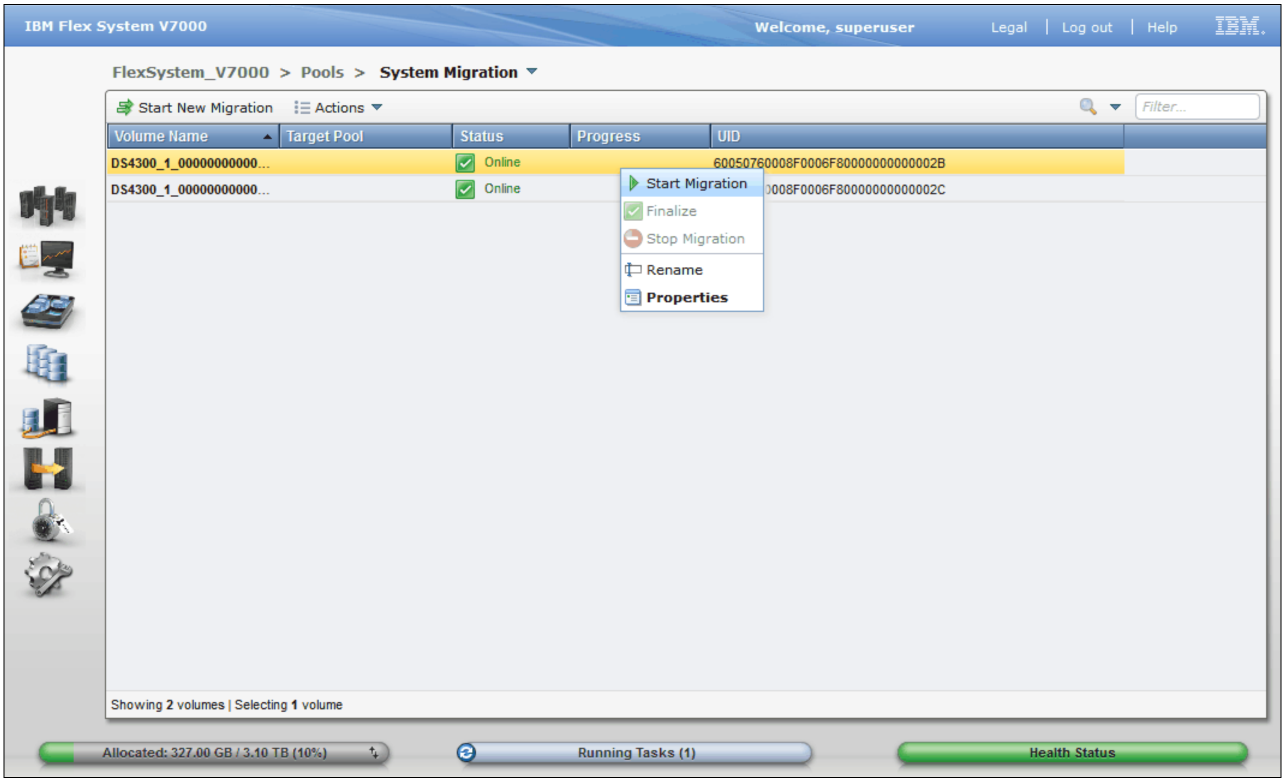


Figure 8-42 Start Migration of Image mode disk

Choose the wanted target storage pool to migrate the data into and click **Add to Pool** (Figure 8-43). Only regular pools, and no migration pools, are offered as a destination.

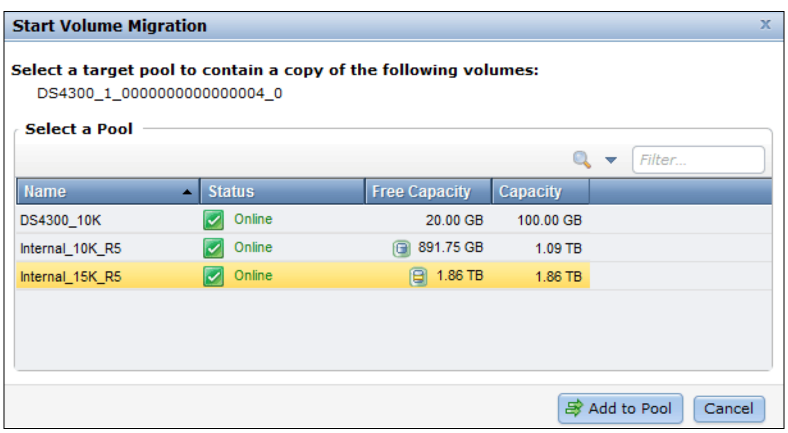


Figure 8-43 Migrate Image Mode Volume into a regular storage pool

The migration itself internally uses the volume copy function, which creates a second copy of the existing volume in the chosen target pool. For more information about the volume copy function, refer to Chapter 9, "IBM Flex System V7000 Storage Node Copy Services" on page 341. If the migration process is interrupted for any reason, select the volume being migrated and click **Actions** → **Stop Migration** (Figure 8-44 on page 327). The process of adding a secondary volume copy is interrupted.

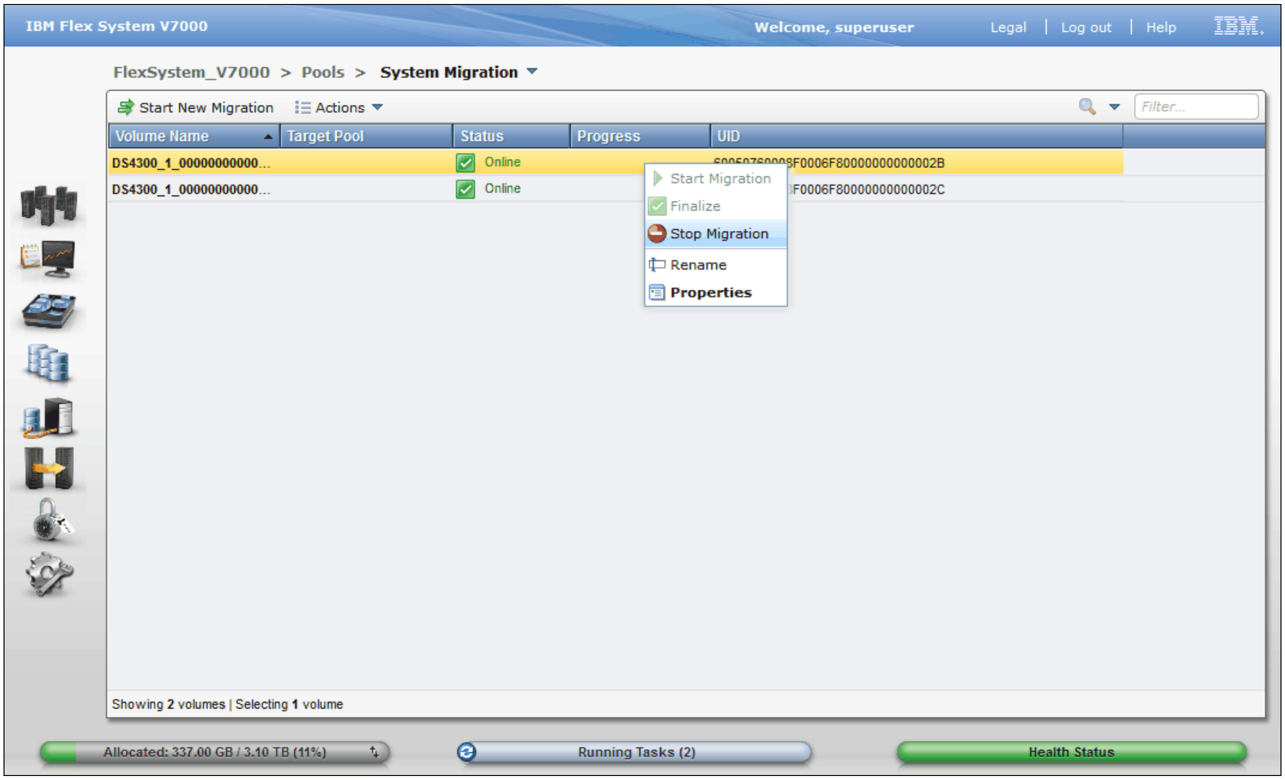


Figure 8-44 Stop Migration

When the migration progress reaches 100%, click **Finalize** to complete the migration process (Figure 8-45).

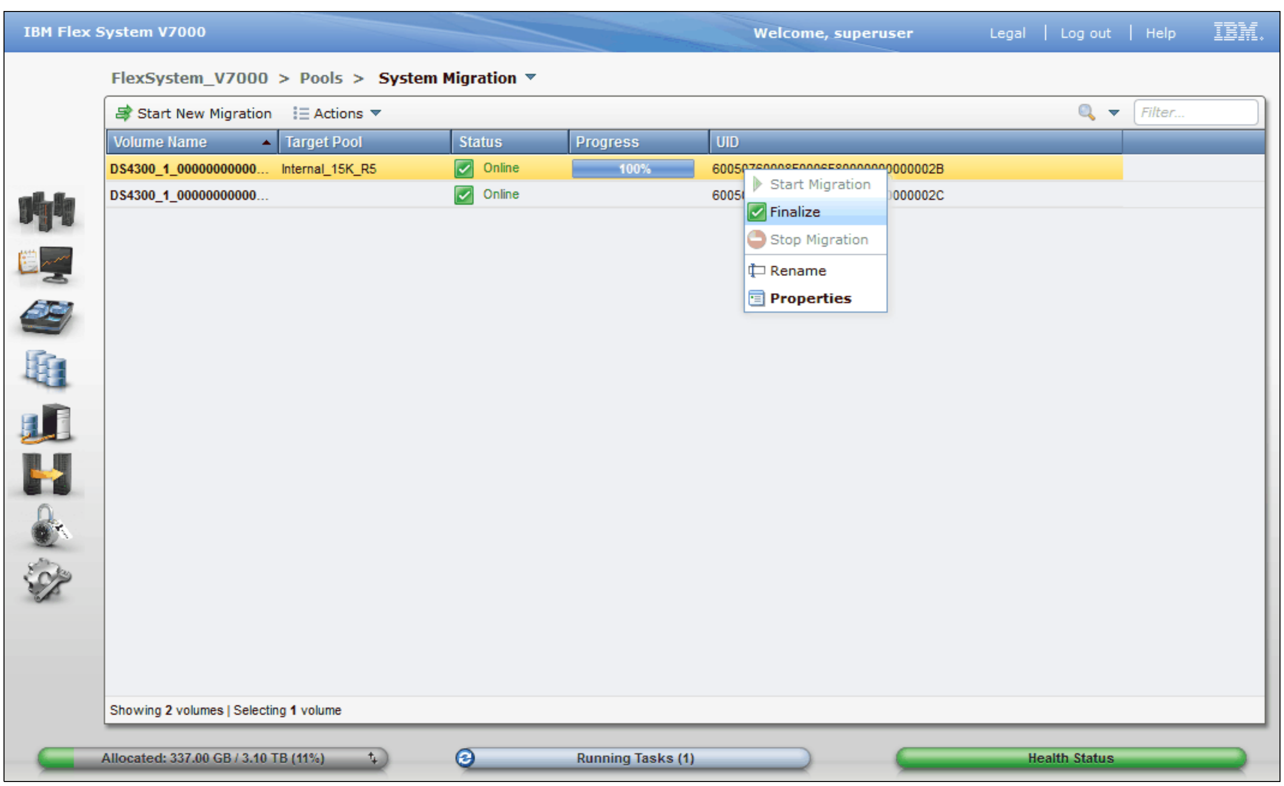


Figure 8-45 Finalize volume migration

The original volume copy on the image mode MDisk is deleted and the newly created copy is kept. The source image mode MDisk is removed from the Migration Pool and changes its mode back to unmanaged.

Figure 8-46 shows that we confirm that the link between the original source volume and the destination volume is finally broken where after the migration finalizes.

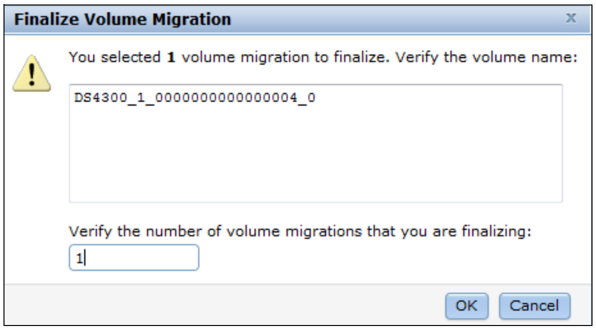


Figure 8-46 Volume Migration Finalizes

At this point the migration of the Image mode disk has finished and the Migration Pool called MigrationPool_8192 will be empty although it has not disappeared. The MDisk that was originally imported will be back as an MDisk that is “Not in a Pool” from where it is now ready for either unmapping from the source disk array or to be used as part of another storage pool by using the “Add to Pool” option. The “Add to Pool” option includes the MDisk in an existing storage pool as a managed disk. By doing that new data will be striped across the disk and its original content is overwritten.

8.2.3 RAID action for MDisks

Internal drives in the IBM Flex System V7000 Storage Node are managed as Array mode MDisks, on which several RAID actions can be performed. Select the appropriate Array MDisk by clicking **Pools** → **MDisks by Pools**, and then click **Actions** → **RAID Actions** (Figure 8-47 on page 329).

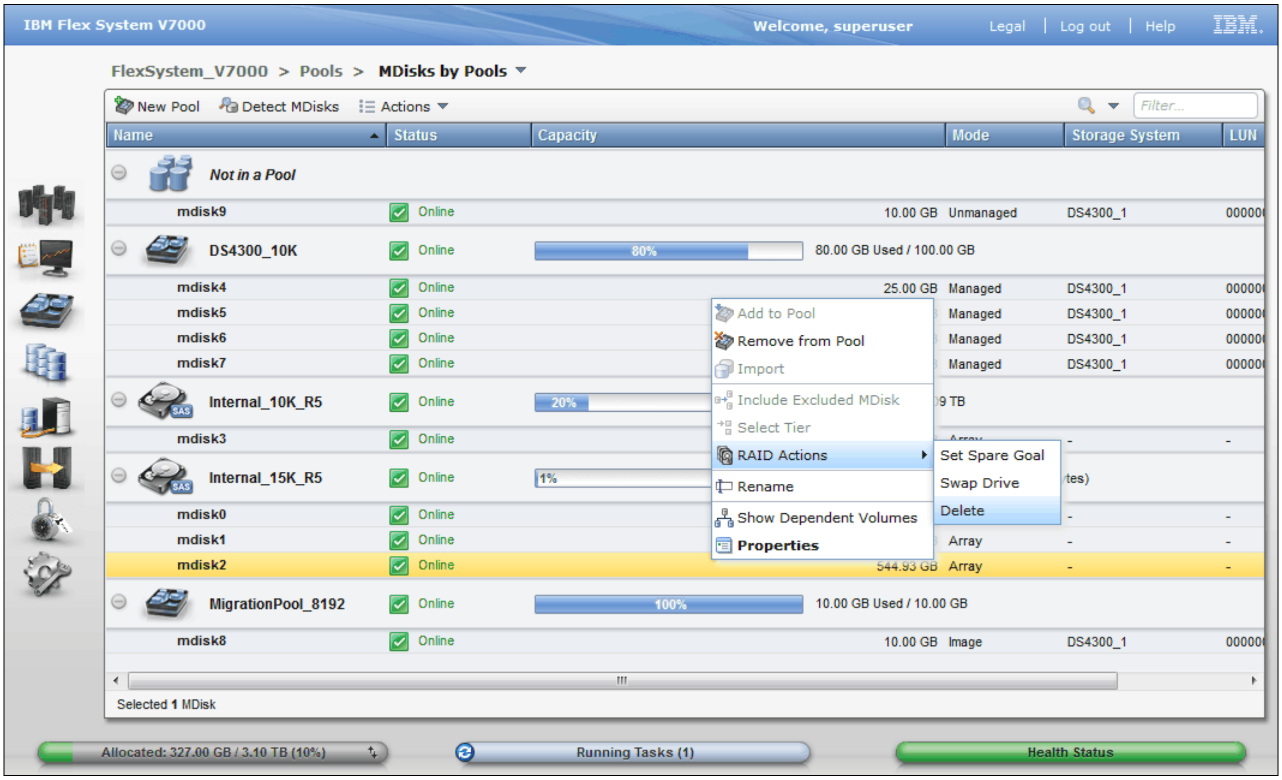


Figure 8-47 RAID actions on Array MDisks

You can choose the following RAID actions:

- **Set Spare Goal**: You can set the wanted number of spare drives that is required to protect the array from drive failures. If the number of drives assigned as Spare does not meet the configured spare goal, an error is logged in the event log that reads “Array MDisk is not protected by sufficient spares”. This error can be fixed by either replacing failed drives, which occupy a spare drive, or by adding additional drives as spare drives. Remember that during the internal drive configuration that automatically spare drives are assigned according to the chosen RAID preset’s spare goals, as described in 8.1.2, “Configuring internal storage” on page 303.
- **Swap drive:** The Swap Drive action can be used to replace a drive in the array with another drive with the status of Candidate or Spare. This action is used to replace a drive that has failed, or is expected to fail soon. For example, as indicated by an error message in the event log. Select an MDisk containing the drive to be replaced and click **RAID Actions** → **Swap Drive**. In the Swap Drive window (Figure 8-48 on page 330), select the member drive to be replaced and click **Next**.

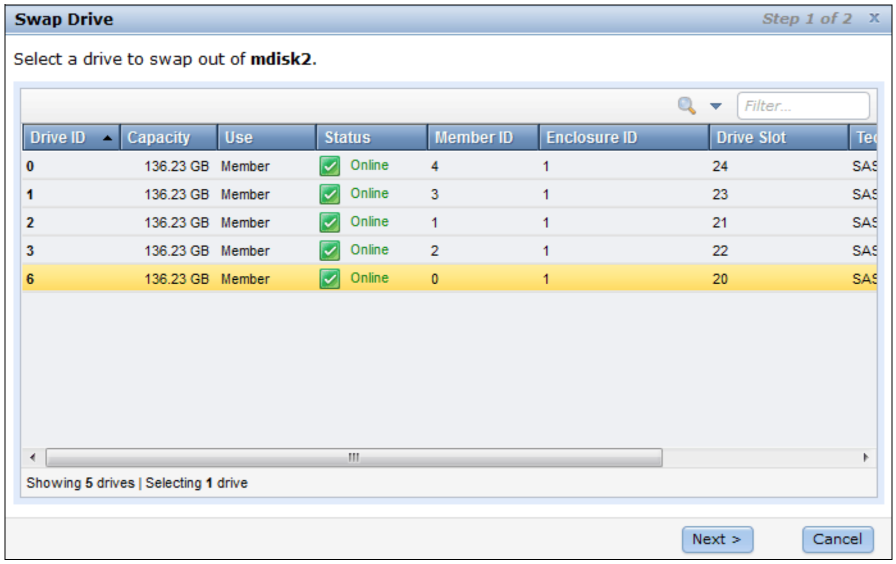


Figure 8-48 Select a drive to swap out of the MDisk

In Figure 8-49, select an available Candidate or Spare drive to swap into the Array MDisk and click **Finish**.

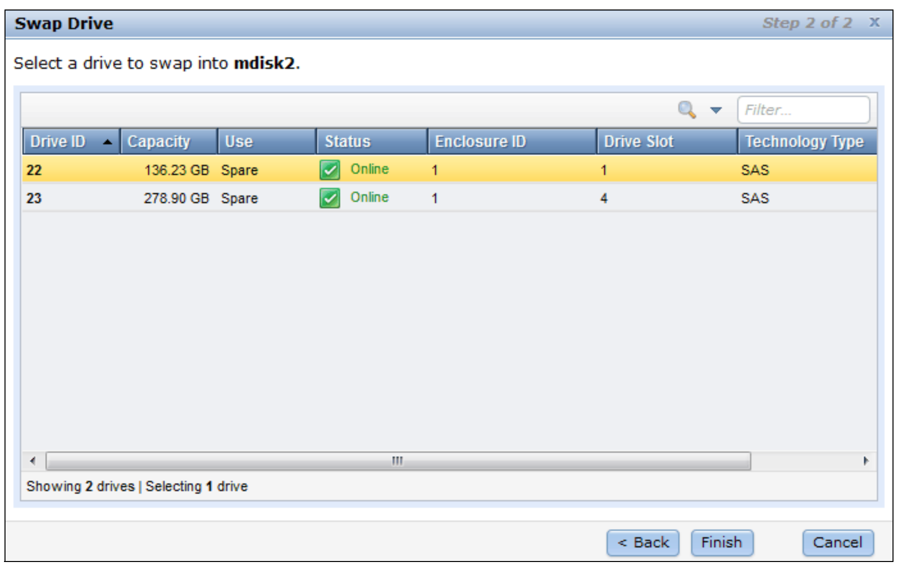


Figure 8-49 Select a drive to swap into the MDisk

The exchange of the drives starts and runs in the background, and the volumes on the affected MDisk remain accessible.

- **Delete**: An Array MDisk can be deleted by clicking **RAID Actions** → **Delete**. A confirmation of the deletion by entering the correct number of MDisks to be deleted is required (Figure 8-50 on page 331). You need to confirm the number of array MDisks you want to delete. If there is data on the MDisk, it can be deleted only by checking the item **Delete the RAID array MDisk even if it has data on it. The system migrates the data to other MDisks in the pool**.

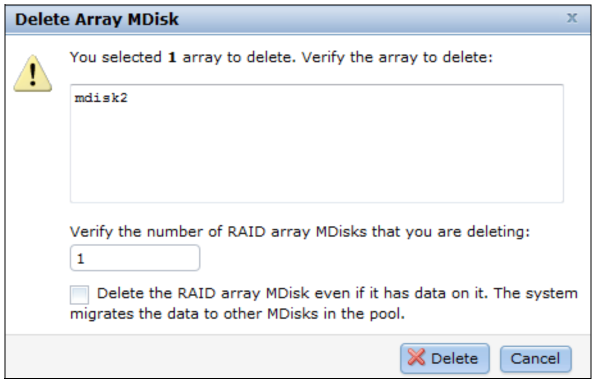


Figure 8-50 Delete Array MDisks window

Available capacity: Make sure that you have enough available capacity left in the storage pool for the data on the MDisks to be removed.

Data on MDisks is migrated to other MDisks in the pool, assuming enough space is available on the remaining MDisks in the pool. After deleting an MDisk from a pool, its member drives return to candidate mode.

8.2.4 Selecting the tier for MDisks

The IBM Flex System V7000 Storage Node Easy Tier feature is described in 2.5.12, "Easy Tier" on page 68. In this section, we show how to adjust the tier settings.

There are two tiers:

- Generic SSD tier for storage made of solid-state drives, which is the faster-performing storage
- Generic HDD tier for everything else

Internal drives have their tier assigned automatically by the IBM Flex System V7000 Storage Node. MDisks on external storage systems are assigned the *generic HDD tier* by default. This setting can be changed manually by the user. To assign a specific tier to an MDisk, click **Pools** → **MDisks by Pool** and click **Select Tier** from the **Actions** drop-down menu (Figure 8-51).

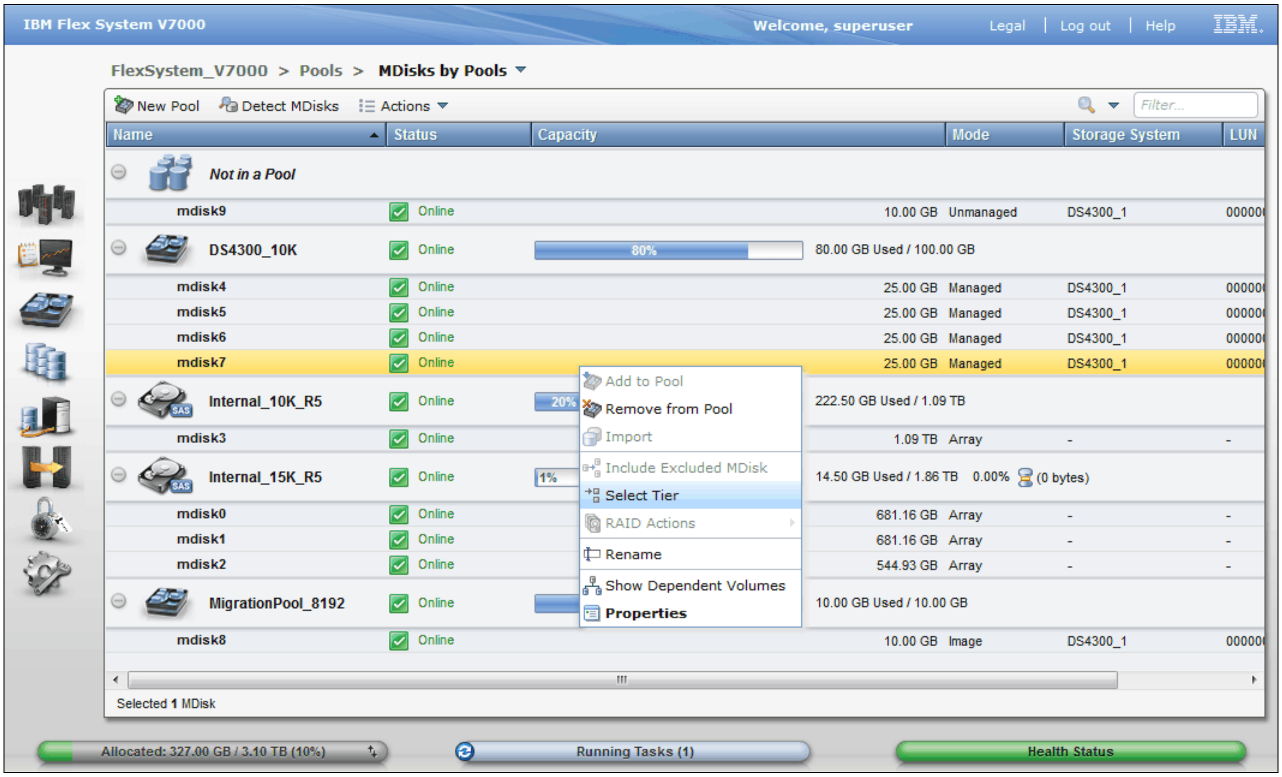


Figure 8-51 Select Tier for an MDisk

For demonstration purposes, we assign the tier **Solid-State Drive** to mdisk7 (Figure 8-52). This MDisk is a LUN made of Fibre Channel hard disk drives in an external storage system. The tier that was assigned by default is Hard Disk Drive.

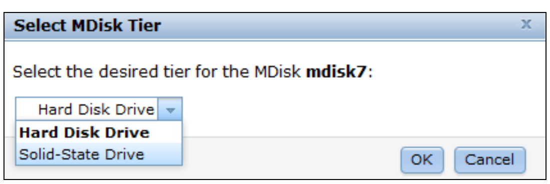


Figure 8-52 Assign desired tier to an MDisk

After the action has been completed successfully, the MDisk can be found in the SSD tier (Figure 8-53).

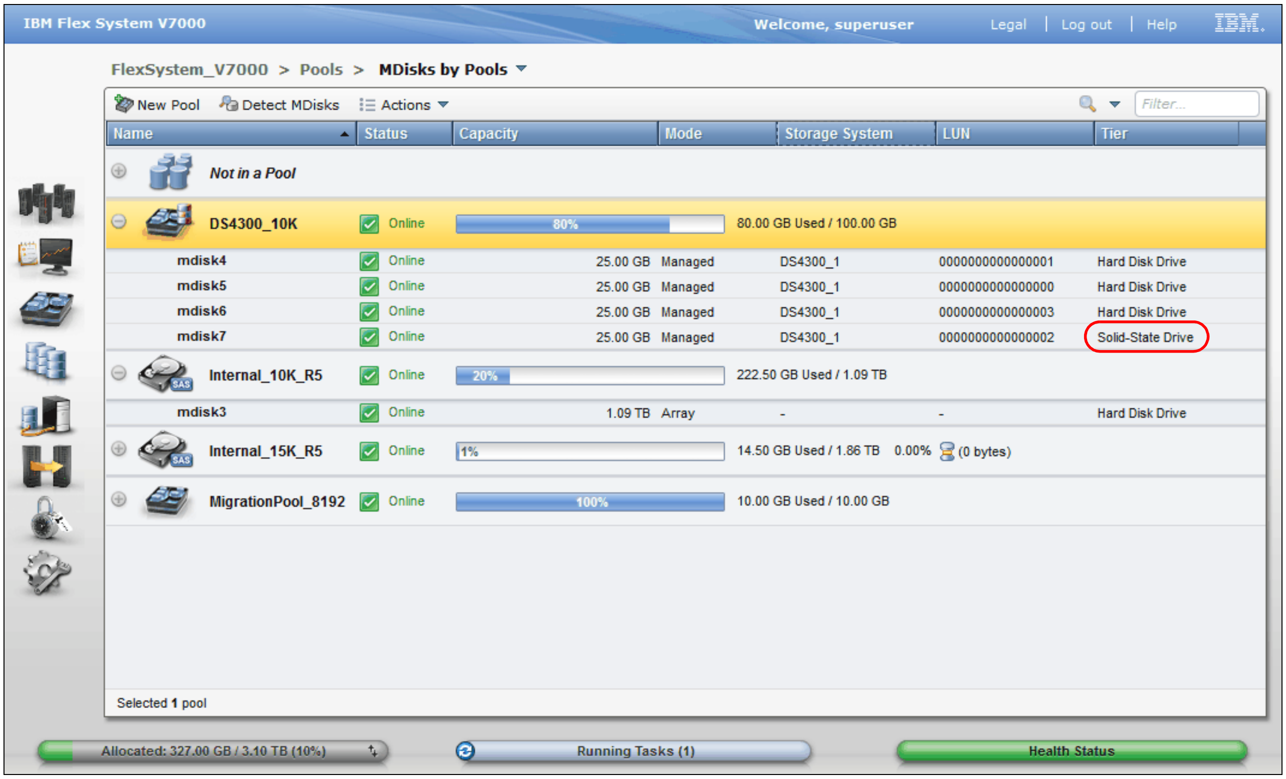


Figure 8-53 The desired tier has been assigned to the MDisk

The action we are performing above instructs the system to use mdisk7 as a faster disk. The IBM Flex System V7000 Storage Node will then move highly utilized storage extents to this MDisk.

8.2.5 Additional actions on MDisks

Here are some of the additional actions that can be performed on MDisks. Such actions can be performed by clicking **Pools** → **MDisks by Pool** → **Actions** or right clicking the selected MDisk.

- **Detect MDisks**: The Detect MDisks button at the upper left of the MDisks by Pools window is useful if you have external storage controllers in your environment. The Detect MDisk action initiates a rescan of the Fibre Channel network. It discovers any new MDisks that have been mapped to the IBM Flex System V7000 Storage Node storage system, and rebalances MDisk access across the available controller device ports. This action also detects any loss of controller port availability, and updates the IBM Flex System V7000 Storage Node configuration to reflect any changes.

  When external storage controllers are added to the IBM Flex System V7000 Storage Node environment, the IBM Flex System V7000 Storage Node automatically discovers the controllers and the LUNs that are presented by those controllers are listed as unmanaged MDisks. However, if you have attached new storage and the IBM Flex System V7000 Storage Node has not detected it, you might need to use the **Detect MDisk** button before the system detects the new LUNs. If the configuration of the external controllers is modified afterward, the IBM Flex System V7000 Storage Node might be unaware of these

configuration changes. Use the **Detect MDisk** button to rescan the Fibre Channel network and update the list of unmanaged MDisks.

> **MDisks detection:** The Detect MDisks action is asynchronous. Although the task appears to be finished, it still might be running in the background.

- **Include Excluded MDisks**: An MDisk could be excluded from the IBM Flex System V7000 Storage Node because of multiple I/O failures. These failures might be caused, for example, by link errors. After a fabric-related problem has been fixed, the excluded disk can be added back into the IBM Flex System V7000 Storage Node by selecting the MDisks and clicking **Include Excluded MDisk** from the **Actions** drop-down menu.
- **Rename**: MDisks can be renamed by selecting the MDisk and clicking **Rename** from the **Actions** drop-down menu. Input the new name of your MDisk and click **Rename** (Figure 8-54).

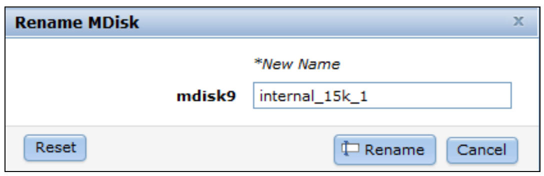


Figure 8-54 Rename MDisk

- **Show Dependent Volumes**: The volumes dependent on a particular MDisk can be displayed by selecting the MDisk and clicking **Show Dependent Volumes** from the **Actions** drop-down list. The volumes are listed with general information (Figure 8-55).

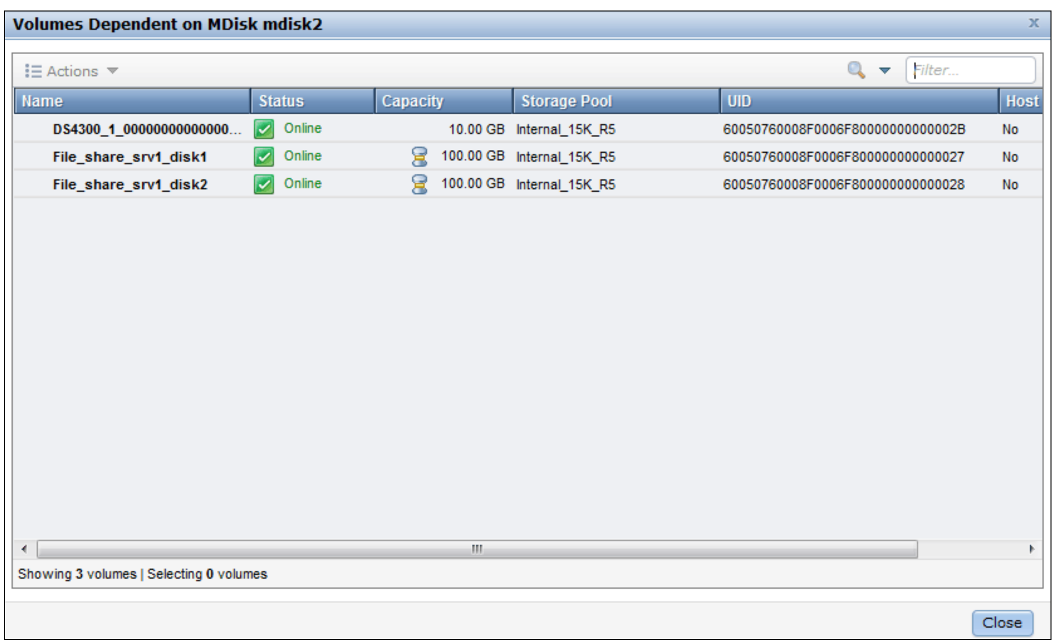


Figure 8-55 Show dependent volumes

Several actions can be taken on volumes by selecting the volume and choosing the action needed from the **Actions** drop-down menu or by right-clicking the volume. For more information about volume actions, refer to Chapter 6, "Basic Volume and Host Configuration" on page 241.

Properties for Mdisks

The Properties action for an MDisk shows the information that you need to identify it. In the MDisks by Pools window, select the MDisk and click **Properties** from the **Actions** drop-down menu. A window opens as shown in Figure 8-56.

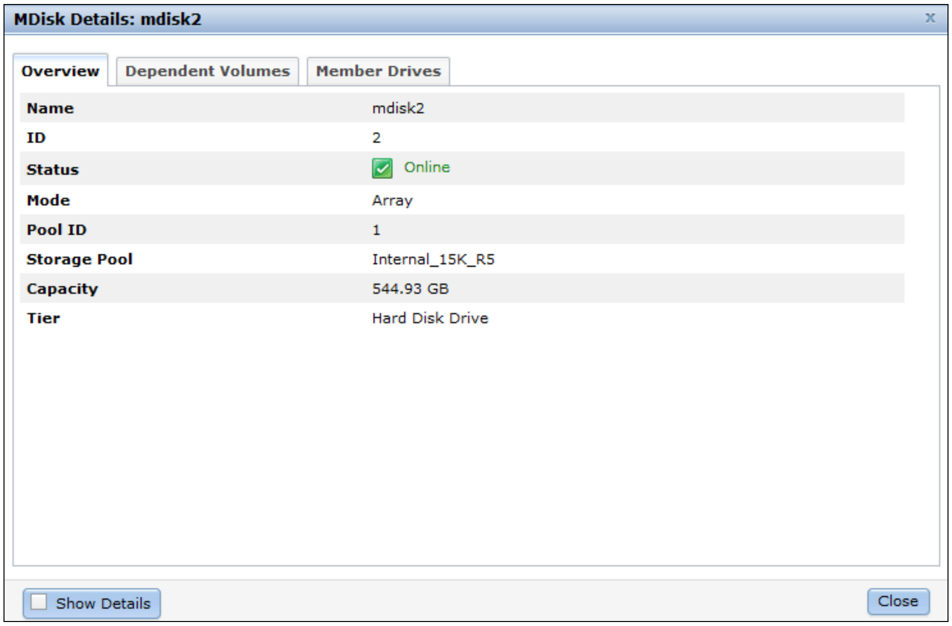


Figure 8-56 MDisk overview with general information

There are three tabs in this information window.

- The Overview tab contains information about the MDisk itself. To show more details, click **Show Details** (Figure 8-57).

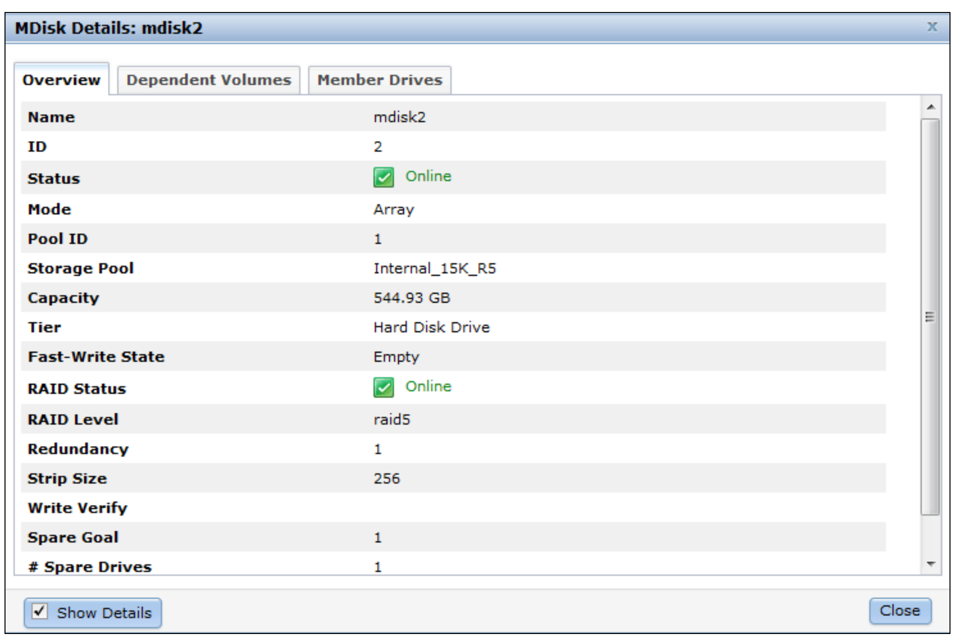


Figure 8-57 MDisk overview with detailed information

- The Member Drives tab is only used with array MDisks. In the Member Drives tab, you find all the member drives of this MDisk (Figure 8-58). Also, all actions described in 8.1.1, “Actions on internal drives” on page 300 can be performed on the drives listed here.

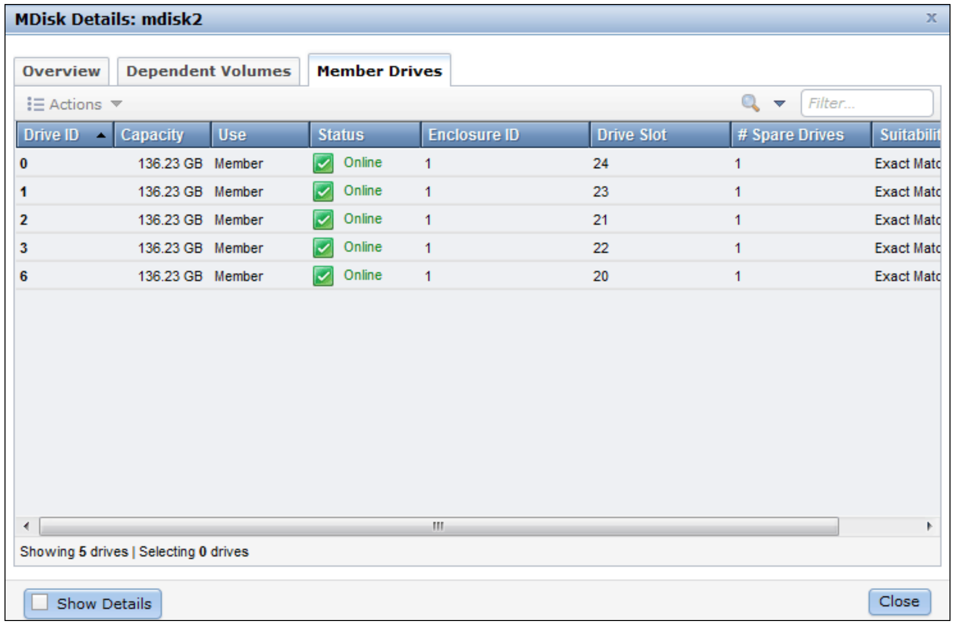


Figure 8-58 Member Drives tab

- The Dependent Volumes tab lists all volumes using extents on this MDisk as shown in Figure 8-59.

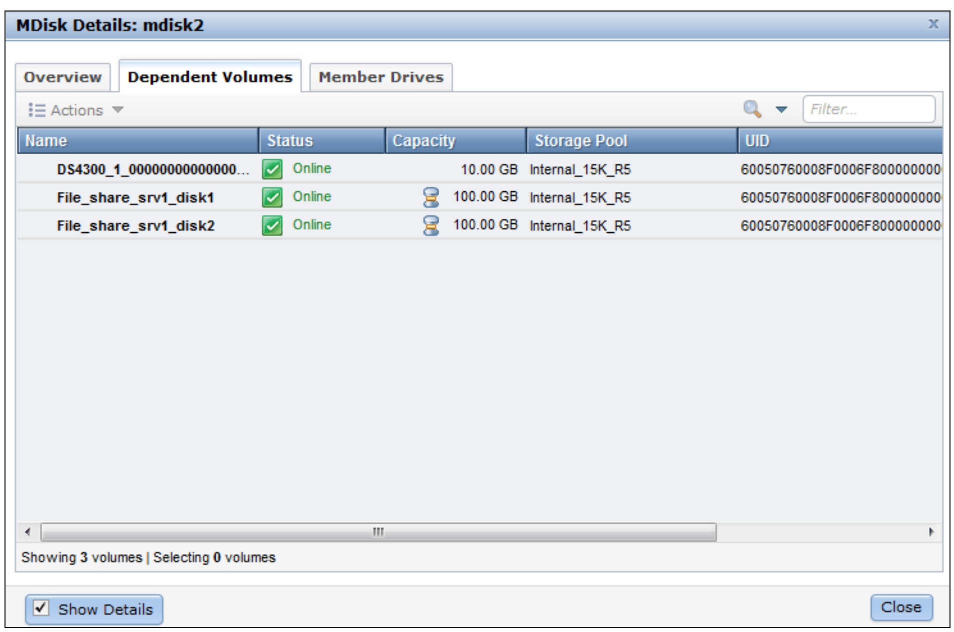


Figure 8-59 Dependent volumes

8.3 Working with storage pools

Storage pools act as a container for MDisks and provision the capacity to volumes. They are comparable to MDisk Groups in IBM SAN Volume Controller or an array or a group of disks managed by a storage controller. IBM Flex System V7000 Storage Node organizes storage in storage pools to ease storage management and make it more efficient. Storage pools and MDisks are managed in MDisks by Pools window. You can access the MDisks by Pools window by clicking **Home** → **Overview** and clicking the **Pools** icon. Extended help information for storage pools is displayed, and if you click **Visit Pools**, the MDisks by Pools window opens. This was shown previously in Figure 8-23 on page 314

The other way to access the Pools window is to click the icon to the left that shows a physical harddisk - the Pools icon.

The MDisk by Pools window (Figure 8-60) allows you to manage storage pools, both internal and as external. All existing storage pools are displayed row-by-row. The first row contains the item Not in a Pool, containing all unmanaged MDisks, if any exist.

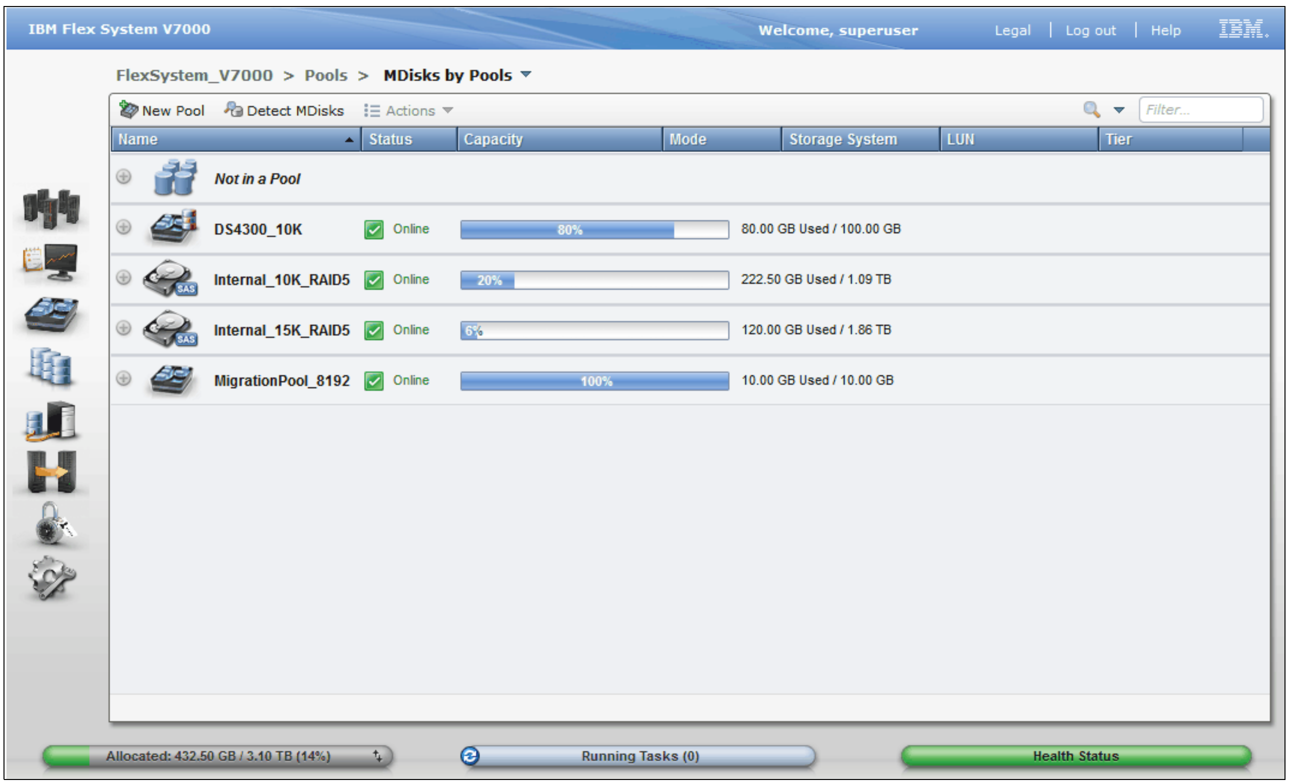


Figure 8-60 MDisks by Pools window

To make the view more clear and structured, filters can be used to show only those storage pools matching the conditions defined in the Filter field in the upper right of the window. In Figure 8-61 on page 338 we show how to filter on *Name* and *Internal* showing us only MDisk Pools with the name *Internal* as part of the name.

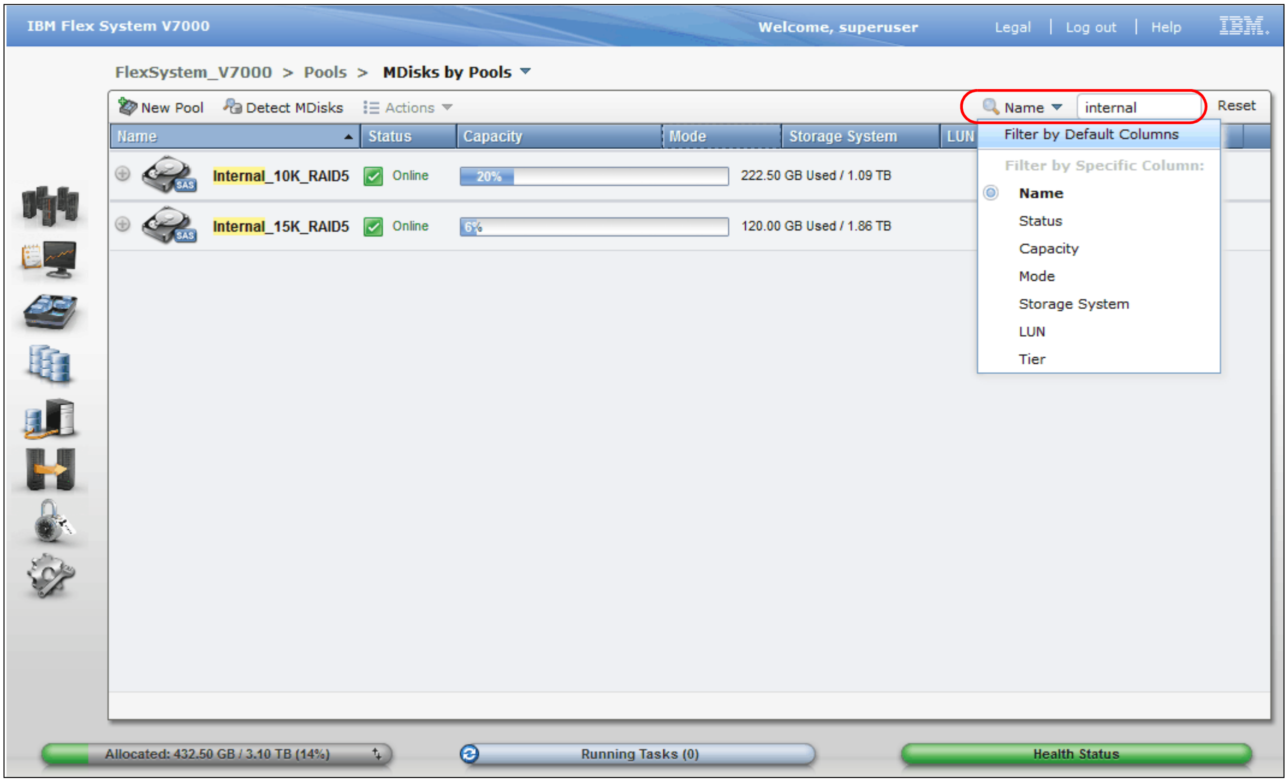


Figure 8-61 Filter the pools view

A few actions can be performed on storage pools using the **Actions** menu. A pool can be *renamed*, its *icon* can be changed, and it can be deleted from here as well.

When you expand a pool's entry by clicking the plus sign (+) to the left of the pool's icon, you get access to the MDisks associated with this pool. You can perform all actions on them, as described in 8.2, "Working with MDisks" on page 313.

Each defined storage pool is displayed along with its assigned icon and *Name*, *Status*, and a graphical indicator showing that the ratio the pool's capacity is allocated to volumes. Also the *Mode* of the *MDisk*, which external storage system it may belong to (if it does), the *LUN ID* and the *Tier* is shown by default.

A red vertical bar inside the indicator determines the warning level at which a Storage pool space warning will be posted in the system's event log. The default warning threshold is at 80% utilization.

To remove a storage pool that is not needed, click **Delete Pool** from the **Actions** drop-down men (Figure 8-62 on page 339).

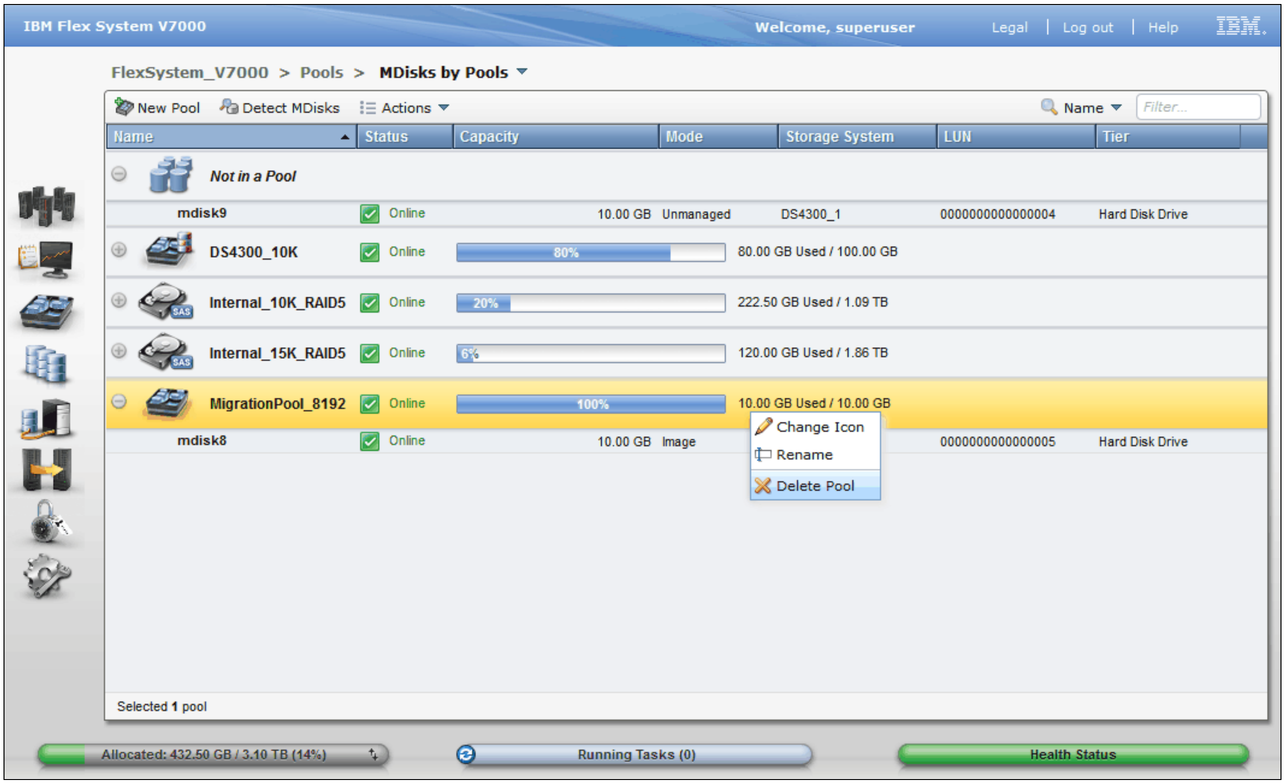


Figure 8-62 Delete Pool window

A confirmation window opens (Figure 8-63). If there are volumes with their only copy on MDisks in the pool to be deleted, they are listed in the confirmation window. If it is safe to delete the volumes and MDisks in this pool, even though they still might be mapped to a host, select the **Delete all volumes, host mappings and MDisks that are associated with this pool** check box and click **Delete**.

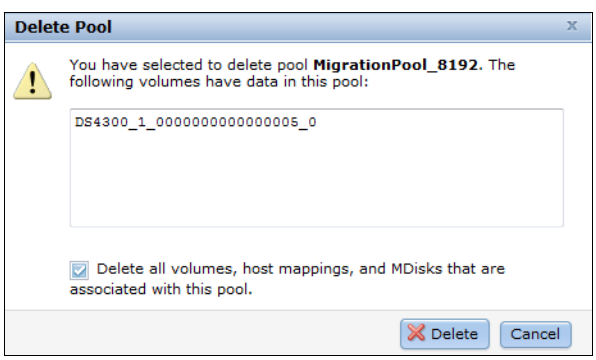


Figure 8-63 Confirm the deletion of the pool

In our example above we are deleting a migration pool that has an *Image Mode* MDisk in it. This means that an external storage system is mapping a LUN to the IBM Flex System V7000 Storage Node, and this LUN is included in the MDisk Pool as an Image of the source disk hence its content is not modified by the IBM Flex System V7000 Storage Node. It can then be mapped to a host and later be migrated to an MDisk Pool with sufficient space. Migrating the Image Mode disk to another Pool makes it managed by the IBM Flex System V7000 Storage Node.

> **Important:** After you delete the pool, all the data in the pool is lost except for the image mode MDisks; their volume definition is deleted, yet the data on the imported MDisk remains untouched.

After you delete a pool, all associated volumes and their host mappings are removed.

Figure 8-64 shows that the MDisks in the Migration Pool is returned to the *Not in a Pool* MDisk group.

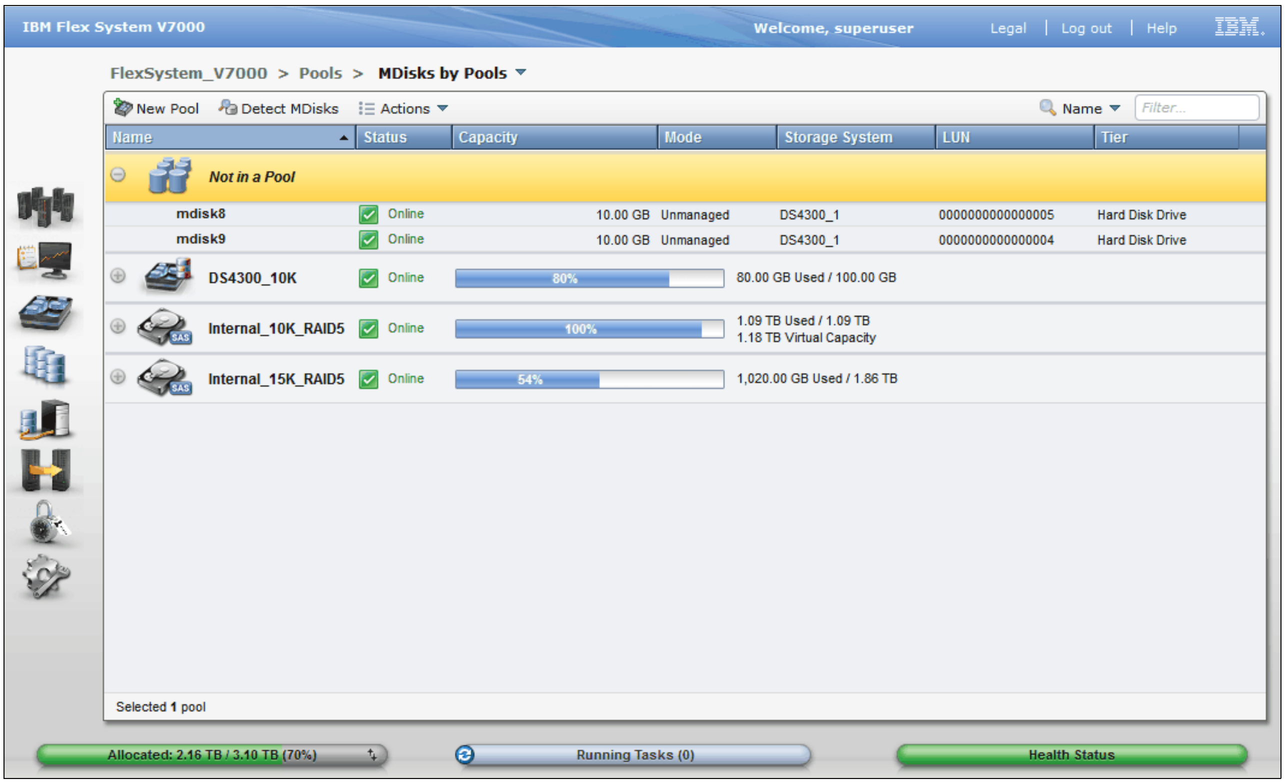


Figure 8-64 The image mode MDisk returns as Not in a Pool

All the managed or image mode MDisks in the pool return to a status of unmanaged after the pool is deleted. If an MDisk pool is deleted, all the array mode MDisks (Internal disks) in the pool are removed and all the member drives return to *Candidate* status.

9

IBM Flex System V7000 Storage Node Copy Services

In this chapter, we describe the Copy Service features provided by the IBM Flex System V7000 Storage Node. All the copy services features of the IBM Flex System V7000 Storage Node are provided by the shared code of the IBM SAN Virtual Controller (SVC) and the IBM Storwize V7000, and all function in the same manner. This commonality allows for the pairing of the functions between these systems for uses such as making data copies, central backups, development, application test, and disaster recovery.

IBM Flex System V7000 Storage Node makes it easy to apply these functions to your environment through its intuitive GUI.

9.1 Services provided

The IBM Flex System V7000 Storage Node features the following copy services capabilities:

- Flashcopy
- Synchronous remote copy
- Asynchronous remote copy

These functions are useful for making data copies for backup, application test, disaster recovery, and so on.

To decide on the functions to be used to meet a need, it is best to plan for the desired results. Two points to be considered for this planning are:

- Recovery point objective (RPO) the point at which the data needs to be able to be recovered to after a disaster occurs.
- Recovery time objective (RTO) the amount of time that the recovery can take to get back up and running after a disaster occurs.

How these objectives are defined can dictate what service capabilities may need to be implemented. in some cases, it may be decided that the use of a combination of the services might better meet the over all need rather than trying to use one specific service. This can be true for cases where a solution may prove to be cost prohibitive for an environment.

9.2 FlashCopy

The FlashCopy function of the IBM Flex System V7000 Storage Node lets you create a *point-in-time copy* of one or more volumes. In this section, we describe the structure of FlashCopy and provide details about its configuration and use.

You can use FlashCopy to solve critical and challenging business needs that require duplication of data of on your source volume. Volumes may remain online and active while you create consistent copies of the data sets. Because the copy is performed at the block level, it operates below the host operating system and cache and is therefore not apparent to the host.

> **Flushing:** Because FlashCopy operates at the block level, below the host operating system and cache, those levels do need to be flushed to create a consistent FlashCopy copy image for a point in time.

While the FlashCopy operation is performed, the source volume is frozen briefly to initialize the FlashCopy bitmap and then I/O is allowed to resume. Although several FlashCopy options require the data to be copied from the source to the target in the background, which can take time to complete, the resulting data on the target volume copy appears to have completed immediately. This task is accomplished through the use of a bitmap (or bit array) that tracks changes to the data after the FlashCopy is initiated, and an indirection layer, which allows data to be read from the source volume transparently.

9.2.1 Business requirements for FlashCopy

When deciding if FlashCopy addresses your needs, you need to adopt a combined business and technical view of the problems you need to solve. Determine what your needs are from a

business perspective, and then determine if FlashCopy fulfills the technical needs of those business requirements.

With an immediately available copy of the data, FlashCopy could be used in various business scenarios, which include:

- Rapidly creating consistent backups of dynamically changing data

  FlashCopy can be used to create backups through periodic execution of FlashCopy; the FlashCopy target volumes can be used to complete a rapid restore of individual files or the entire volume through Reverse FlashCopy (using the `-restore` option).

  The target volumes created by FlashCopy may be also used for backup to tape, by attaching them to another server, which to a great extent relieves the production server. After the copy to tape completes, the target volumes can be discarded if required or kept as a rapid restore copy of the data.

- Rapidly creating consistent copies of production data to facilitate data movement or migration between hosts.

  FlashCopy can be used to facilitate the movement or migration of data between hosts while minimizing downtime for applications. FlashCopy allows application data to be copied from source volumes to new target volumes while applications remain online. After the volumes are fully copied and synchronized, the application can be stopped and then immediately started on the new server accessing the new FlashCopy target volumes. This mode of migration is swifter than other migration methods available through the IBM Flex System V7000 Storage Node, because the size or the speed of the migration is not as limited.

- Rapidly creating copies of production data sets for application development and testing.

  Under normal circumstances, to perform application development and testing, data must be restored from traditional backup media, such as tape. Depending the amount of data and the technology in use, this process could easily take a day or more. With FlashCopy, a copy can be created and online for use in just a few minutes. The time varies based on the application and the data set size.

- Rapidly creating copies of production data sets for auditing purposes and data mining.

  Auditing or data mining normally require the usage of the production applications. This situation can cause high loads for databases track inventories or similar data. With FlashCopy, you can create copies for your reporting and data mining activities. This situation reduces the load on your production systems, increasing their performance.

- Rapidly creating copies of production data sets for quality assurance.

  Quality assurance is an interesting case for FlashCopy. Because traditional methods involve so much time and labor, the refresh cycle is typically extended. This reduction in time required allows much more frequent refreshes of the quality assurance database.

9.2.2 FlashCopy functional overview

FlashCopy occurs between a source volume and a target volume. The source and target volumes must be the same size. Multiple FlashCopy mappings (source-to-target relationships) can be defined, and point-in-time consistency can be maintained across multiple point-in-time mappings using consistency groups. FlashCopy consistency groups are further described in “FlashCopy consistency groups” on page 348.

The minimum granularity that the IBM Flex System V7000 Storage Node supports for FlashCopy is an entire volume; it is not possible to use FlashCopy to copy only part of a

volume. Additionally, the source and target volumes must belong to the same IBM Flex System V7000 Storage Node, but they do not have to reside in the same storage pool.

Before you start a FlashCopy (regardless of the type and options specified) you must run either **prestartfcmap** or **prestartfcconsistgrp**, which puts the con troll canister Cache into write-through mode, which results in the flushing of all I/O currently bound for your source volume. After FlashCopy is started, an effective copy of a source volume to a target volume has been created. The content of the source volume is immediately presented on the target volume and the original content of the target volume is lost. This FlashCopy operation is also referred to as a *time-zero copy* (T_0).

Immediately following the FlashCopy operation, both the source and target volumes are available for use. The FlashCopy operation creates a bitmap that is referenced and maintained to direct I/O requests within the source and target relationship. This bitmap is updated to reflect the active block locations as data is copied in the background from the source to target and updates are made to the source.

Figure 9-1 shows the general process for how FlashCopy works while the full image copy is being completed in the background. Also the handling of the redirection of the host I/O which is being written to the source volume with respect to a T_0 point in time while the target volume is held true to T_0.

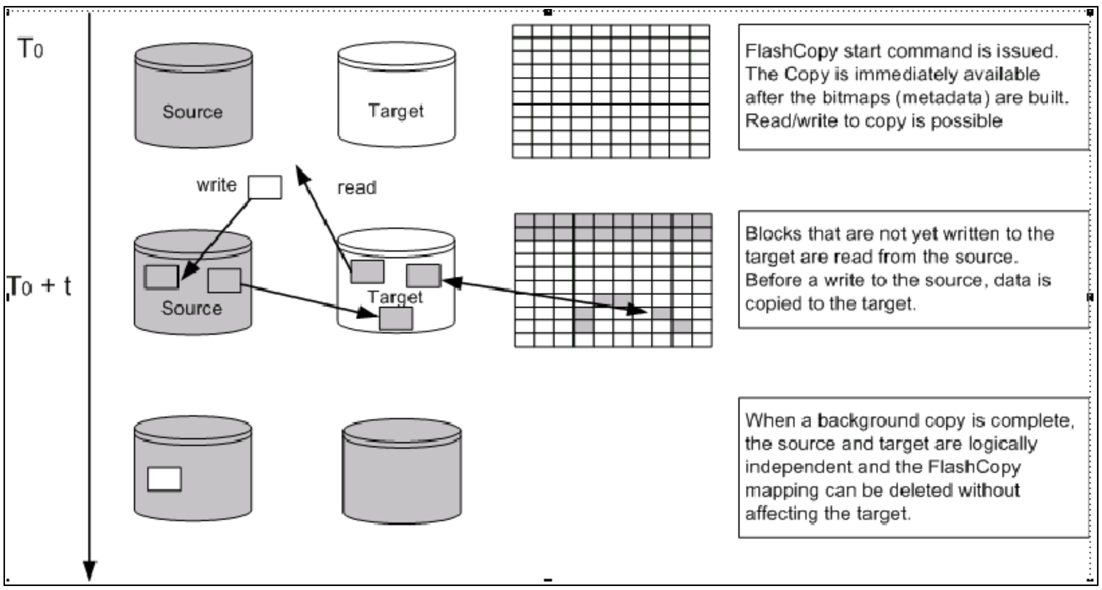


Figure 9-1 Flashcopy processing and redirection of host write I/O

When data is copied between volumes, it is copied in units of address space known as *grains*. Grains are units of data grouped together to optimize the use of the bitmap that track changes to the data between the source and target volume. You have the option of using 64 KB or 256 KB grain sizes; 256 KB is the default. The FlashCopy bitmap contains 1 bit for each grain and is used to track whether the source grain has been copied to the target. The 64 KB grain size consumes bitmap space at a rate of four times the default 256 KB size.

The FlashCopy bitmap dictates read and write behavior for both the source and target volumes as follows:

- Read I/O request to source: Reads are performed from the source volume, which is the same as for non-FlashCopy volumes.
- Write I/O request to source: Writes to the source cause the grains to be copied to the target if it has not already been copied, and then the write is performed to the source.

- Read I/O request to target: Reads are performed from the target if the grains have already been copied; otherwise, the read is performed from the source.
- Write I/O request to target: Writes to the target cause the grain to be copied from the source to the target unless the entire grain is being written, and then the write completes to the target.

FlashCopy mappings

FlashCopy bit mappings define the relationship between a source volume and a target volume. FlashCopy mappings can be either a stand-alone mapping or a member of a consistency group, as described in “FlashCopy consistency groups” on page 348.

Incremental FlashCopy mappings

In an incremental FlashCopy, the initial mapping copies all of the data from the source volume to the target volume. Subsequent FlashCopy mappings only copy data that has been modified since the initial FlashCopy mapping. This action reduces the amount of time that it takes to re-create an independent FlashCopy image. You can define a FlashCopy mapping as incremental only when you create the FlashCopy mapping.

Multiple target FlashCopy mappings

You can copy up to 256 target volumes from a single source volume. Each relationship between a source and target volume is managed by a unique mapping such that a single volume can be the source volume for up to 256 mappings.

Each of the mappings from a single source can be started and stopped independently. If multiple mappings from the same source are active (in the copying or stopping states), a dependency exists between these mappings.

If a single source volume has multiple target FlashCopy volumes, the write to the source volume does not cause its data to be copied to all of the targets. Instead, it is copied to the newest target volume only. The older targets refer to new targets first before referring to the source. A dependency relationship exists between a particular target and all newer targets that share a source until all data has been copied to this target and all older targets.

Cascaded FlashCopy mappings

The cascaded FlashCopy function allows a FlashCopy target volume to be the source volume of another FlashCopy mapping. Up to 256 mappings can exist in a cascade. If cascaded mappings and multiple target mappings are used, a tree of up to 256 mappings can be created.

Cascaded mappings differ from multiple target FlashCopy mappings in depth. Cascaded mappings have an association in the manner of A > B > C, while multiple target FlashCopy has an association in the manner A > B1 and A > B2.

Background copy rate

The background copy rate is a property of a FlashCopy mapping defined as a value of 0 - 100. The background copy rate can be defined and dynamically changed for individual FlashCopy mappings. A value of 0 disables background copy. This option is also called the *no-copy option*, which provides pointer-based images for limited lifetime uses.

With FlashCopy background copy, the source volume data is copied to the corresponding target volume in the FlashCopy mapping. If the background copy rate is set to 0, which means disable the FlashCopy background copy, only data that changed on the source volume is copied to the target volume. The benefit of using a FlashCopy mapping with background copy enabled is that the target volume becomes a real independent clone of the FlashCopy

mapping source volume after the copy is complete. When the background copy is disabled, the target volume only remains a valid copy of the source data while the FlashCopy mapping remains in place. Copying only the changed data saves your storage capacity (assuming it is thin provisioned and `-rsize` has been correctly setup.)

The relationship of the background copy rate value to the amount of data copied per second is shown in Table 9-1.

Table 9-1 Background copy rate

Copy / Cleaning Rate Values	Data copied per second	Grains per second (256 KB grain)	Grains per second (64 KB grain)
1 - 10	128 KB	0.5	2
11 - 20	256 KB	1	4
21 - 30	512 KB	2	8
31 - 40	1 MB	4	16
41 - 50	2 MB	8	32
51 - 60	4 MB	16	64
61 - 70	8 MB	32	128
71 - 80	16 MB	64	256
81 - 90	32 MB	128	512
91 - 100	64 MB	256	1024

Data copy rate: The data copy rate remains the same regardless of the FlashCopy grain size. The difference is the number of grains copied per second. The grain size can be either 64 or 256 KB. The smaller size consumes more bitmap space and thus limits the total amount of FlashCopy space possible, but may be more efficient regarding the amount of data moved, depending on your environment.

Cleaning rate

The cleaning rate provides a method for FlashCopy copies with dependant mappings (either multiple targets or cascaded) to be able to complete their background copies before their source goes offline or is deleted after a stop has been issued.

When you create or modify a FlashCopy mapping, you can specify a cleaning rate for the FlashCopy mapping that is independent of the background copy rate. The cleaning rate is also defined as a value of 0 - 100, which has the same relationship to data copied per second as the background copy rate (Table 9-1).

The cleaning rates controls the rate at which the cleaning process operates. The purpose of the cleaning process is to copy (or flush) data from FlashCopy source volumes upon which there are dependent mappings to their targets. For cascaded and multiple target FlashCopy, the source maybe a target for another FlashCopy target or even a source for a chain (cascade) of FlashCopy mappings. The cleaning process must complete before the FlashCopy mapping can go to the stopped state. This feature is the distinction between *stopping* and *stopped* states which prevents data access interruption for dependent mappings, when their source is issued a stop.

FlashCopy mapping states

At any point in time, a mapping is in one of the following states:

- Idle or Copied

  The source and target volumes act as independent volumes even if a mapping exists between the two. Read and write caching is enabled for both the source and the target volumes.

  If the mapping is incremental and the background copy is complete, the mapping only records the differences between the source and target volumes. If the connection to both nodes in the IBM Storwize V7000 storage system that the mapping is assigned to is lost, the source and target volumes go offline.

- Copying

  The copy is in progress. Read and write caching is enabled on the source and the target volumes.

- Prepared

  The mapping is ready to start. The target volume is online, but is not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the SCSI front end as a hardware error. If the mapping is incremental and a previous mapping has completed, the mapping only records the differences between the source and target volumes. If the connection to both nodes in the IBM Storwize V7000 storage system that the mapping is assigned to is lost, the source and target volumes go offline.

- Preparing

  The target volume is online, but not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the SCSI front end as a hardware error. Any changed write data for the source volume is flushed from the cache. Any read or write data for the target volume is discarded from the cache. If the mapping is incremental and a previous mapping has completed, the mapping records only the differences between the source and target volumes. If the connection to both nodes in the IBM Storwize V7000 storage system that the mapping is assigned to is lost, the source and target volumes go offline.

- Stopped

  The mapping is stopped because either you issued a stop command or an I/O error occurred. The target volume is offline and its data is lost. To access the target volume, you must restart or delete the mapping. The source volume is accessible and the read and write cache is enabled. If the mapping is incremental, the mapping is recording write operations to the source volume. If the connection to both nodes in the IBM Storwize V7000 storage system that the mapping is assigned to is lost, the source and target volumes go offline.

- Stopping

  The mapping is in the process of copying data to another mapping. If the background copy process is complete, the target volume is online while the stopping copy process completes. If the background copy process is not complete, data is discarded from the target volume cache. The target volume is offline while the stopping copy process runs. The source volume is accessible for I/O operations.

- Suspended

  The mapping started, but it did not complete. Access to the metadata is lost, which causes both the source and target volume to go offline. When access to the metadata is restored, the mapping returns to the copying or stopping state and the source and target volumes return online. The background copy process resumes.

Any data that has not been flushed and has been written to the source or target volume before the suspension is in cache until the mapping leaves the suspended state.

FlashCopy consistency groups

Consistency groups address the requirement to preserve point-in-time data consistency across multiple volumes for applications having related data that spans them. For these volumes, consistency groups maintain the integrity of the FlashCopy by ensuring that *dependent writes*, which are described in more detail in "Dependent writes" on page 348, are executed in the application's intended sequence.

When consistency groups are used, the FlashCopy commands are issued to the FlashCopy consistency group, which performs the operation on all FlashCopy mappings contained within the consistency group.

Figure 9-2 shows a consistency group consisting of two FlashCopy mappings.

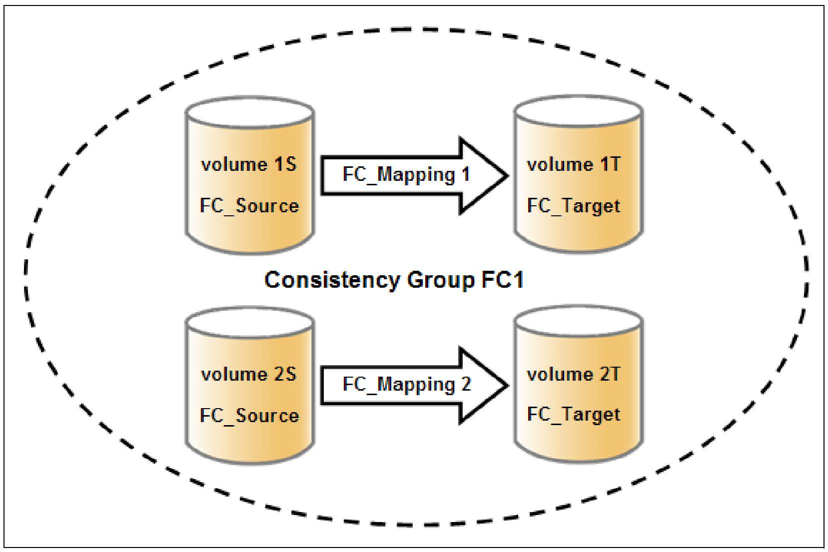


Figure 9-2 FlashCopy consistency group

> **FlashCopy mapping management:** After an individual FlashCopy mapping has been added to a consistency group, it can only be managed as part of the group; operations such as start and stop are no longer allowed on the individual mapping.

Dependent writes

To illustrate why it is crucial to use consistency groups when a data set for a database application spans across multiple volumes, consider the following typical sequence of writes for the database to perform an update transaction:

1. A write is executed to update the database log, indicating that a database update is about to be performed.
2. A second write is executed to complete the actual update to the database.
3. A third write is executed to update the database log, indicating that the database update has successfully completed.

The database ensures the correct ordering of these writes by waiting for each step to complete before starting the next step. However, if the database log (updates 1 and 3) and the database itself (update 2) are on separate volumes, when creating FlashCopy images it is possible for the FlashCopy of the database volume to occur before the FlashCopy of the database log. This situation can result in the target volumes seeing writes (1) and (3) but not (2), because the FlashCopy of the database volume occurred before the write was completed.

In this scenario, if the database was restarted using the backup that was made from the FlashCopy target volumes, the database log indicates that the transaction had completed successfully when in fact it had not. This situation occurs because the FlashCopy of the volume with the database file was started (bitmap was created) and the write updates (1 and 3) to the logs completed before the write to the database (2) had completed to it's volume in turn. Therefore, the transaction is lost and the integrity of the database is in question.

To overcome the issue of dependent writes across volumes and to create a consistent image of the client data, it is necessary to perform a FlashCopy operation on multiple volumes in order of operation using consistency groups.

With the use of consistency groups the action of stopping and flushing the database volumes would all occur simultaneously so that all three of the write operations would be flushed to their volumes as required to make the database transaction complete and the FlashCopy image across all the volumes would be consistent.

A FlashCopy consistency group can contain up to 512 FlashCopy mappings. The more mappings you have, the more time it takes to prepare the consistency group. FlashCopy commands can then be issued to the FlashCopy consistency group and simultaneously for all of the FlashCopy mappings that are defined in the consistency group. For example, when starting the FlashCopy for the consistency group, all FlashCopy mappings in the consistency group are started at the same time, resulting in a point-in-time copy that is consistent across all FlashCopy mappings that are contained in the consistency group.

A consistency group aggregates FlashCopy mappings, not volumes. Thus, where a source volume has multiple FlashCopy mappings, they can be in the same or separate consistency groups. If a particular volume is the source volume for multiple FlashCopy mappings, you might want to create separate consistency groups to separate each mapping of the same source volume. Regardless of whether the source volume with multiple target volumes is in the same consistency group or in separate consistency groups, the resulting FlashCopy produces multiple identical copies of the source data.

The consistency group can be specified when the mapping is created. You can also add the FlashCopy mapping to a consistency group or change the consistency group of a FlashCopy mapping later. Do not place stand-alone mappings into a consistency group, because they become controlled as part of that consistency group.

FlashCopy consistency group states

At any point in time, a FlashCopy consistency group is in one of the following states:

- Idle or Copied

  All FlashCopy Mappings in this consistency group are in the Idle or Copied state.

- Preparing

  At least one FlashCopy mapping in this consistency group is in the Preparing state.

- Prepared

The consistency group is ready to start. While in this state, the target volumes of all FlashCopy mappings in this consistency group are not accessible.

- Copying

  At least one FlashCopy mapping in the consistency group is in the Copying state and no FlashCopy mappings are in the Suspended state.

- Stopping

  At least one FlashCopy mapping in the consistency group is in the Stopping state and no FlashCopy mappings are in the Copying or Suspended state.

- Stopped

  The consistency group is stopped because either you issued a command or an I/O error occurred.

- Suspended

  At least one FlashCopy mapping in the consistency group is in the Suspended state.

- Empty

  The consistency group does not have any FlashCopy mappings.

Reverse FlashCopy

Reverse FlashCopy enables FlashCopy targets to become restore points for the source without breaking the FlashCopy relationship and without waiting for the original copy operation to complete. It supports multiple targets and multiple rollback points.

A key advantage of Reverse FlashCopy is that it does not delete the original target, thus allowing processes using the target, such as a tape backup, to continue uninterrupted.

You can also create an optional copy of the source volume that is made before starting the reverse copy operation. This copy restores the original source data, which can be useful for diagnostic purposes.

Figure 9-3 shows an example of a reverse FlashCopy scenario.

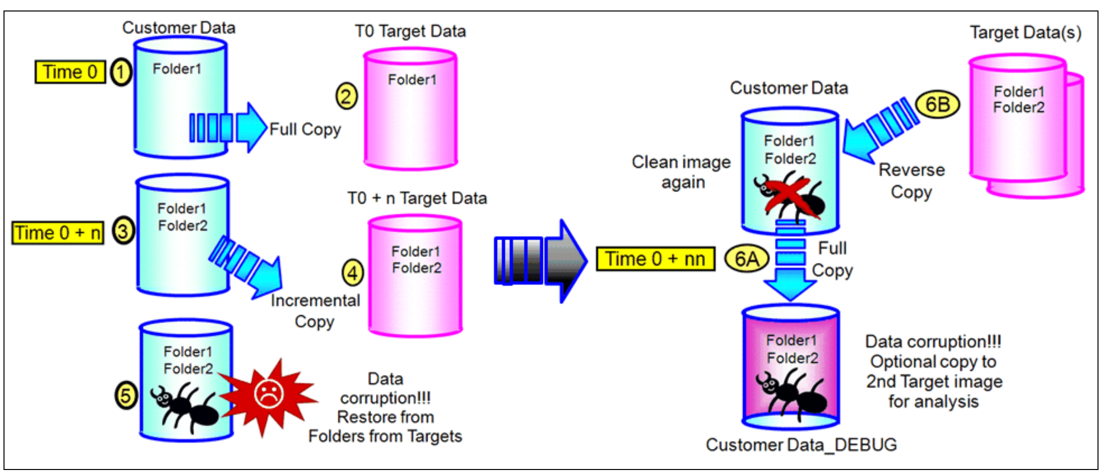


Figure 9-3 Reverse FlashCopy scenario

To restore from an FlashCopy backup, complete the following steps:

1. Optionally, create a target volume (6A) for Customer Data_DEBUG for later problem analysis prior to the Reverse Copy (6B) from previous target volume (Target Data) of the time desired (T0 or T0 + n) to the production volume (Customer Data).

2. Create a FlashCopy mapping for the backup to be restored with the target volume being the source volume and the original source production as the target volume, if this mapping does not already exist.
3. Start the FlashCopy map (Target Data -> Customer Data) with the `-restore` option to copy the backup data onto the production disk.

> **The -restore option:** In the GUI, the `-restore` option is applied automatically when you start the FlashCopy mapping from Target Data to Customer Data; when using the CLI, you need to add the `-restore` option to the command manually.

The production disk is instantly available with the backup data.

Regardless of whether the initial FlashCopy map (Customer Data to Target Data) is incremental, the Reverse FlashCopy operation only copies the modified data.

Consistency groups are reversed by creating a set of new “reverse” FlashCopy maps and adding them to a new “reverse” consistency group. Consistency groups cannot contain more than one FlashCopy map with the same target volume.

9.2.3 Planning for FlashCopy

There are several items that must be considered before performing a FlashCopy, which we describe in this section.

Guidelines for FlashCopy implementation

Consider the following guidelines for FlashCopy implementation:

- The source and target volumes must be on the same IBM Storwize V7000 storage system.
- The source and target volumes do not need to be in the same storage pool.
- The FlashCopy source and target volumes can be thin-provisioned.
- The source and target volumes must be the same size. The size of the source and target volumes cannot be altered (increased or decreased) while a FlashCopy mapping is defined.
- FlashCopy operations perform in direct proportion to the performance of the source and target disks. If you have a fast source disk and slow target disk, the performance of the source disk is reduced because it must wait for the write operation to occur at the target before it can write to the source.

Maximum configurations for FlashCopy

Table 9-2 shows some of the FlashCopy maximum configurations.

Table 9-2 FlashCopy maximum configurations

FlashCopy property	Maximum
FlashCopy targets per source	256
FlashCopy mappings per cluster	4,096
FlashCopy consistency groups per cluster	127

FlashCopy property	Maximum
FlashCopy mappings per consistency group	512

FlashCopy presets

The IBM Storwize V7000 storage system provides three FlashCopy presets, named Snapshot, Clone, and Backup, to simplify the more common FlashCopy operations (Table 9-3).

Table 9-3 FlashCopy presets

Preset	Purpose
Snapshot	Creates a point-in-time view of the production data. The snapshot is not intended to be an independent copy, but is used to maintain a view of the production data at the time the snapshot is created. This preset automatically creates a thin-provisioned target volume with 0% of the capacity allocated at the time of creation. The preset uses a FlashCopy mapping with 0% background copy so that only data written to the source or target is copied to the target volume.
Clone	Creates an exact replica of the volume, which can be changed without affecting the original volume. After the copy operation completes, the mapping that was created by the preset is automatically deleted. This preset automatically creates a volume with the same properties as the source volume and creates a FlashCopy mapping with a background copy rate of 50. The FlashCopy mapping is configured to automatically delete itself when the FlashCopy mapping reaches 100% completion.
Backup	Creates a point-in-time replica of the production data. After the copy completes, the backup view can be refreshed from the production data, with minimal copying of data from the production volume to the backup volume. This preset automatically creates a volume with the same properties as the source volume. The preset creates an incremental FlashCopy mapping with a background copy rate of 50.

Presets: All of the presets can be adjusted by using the Advanced Settings expandable section in the GUI.

9.2.4 Managing FlashCopy using the GUI

The IBM Storwize V7000 storage system provides a separate function icon to access copy service management. There are three windows for managing FlashCopy under the Copy Services function icon, including:

- FlashCopy
- Consistency Groups
- FlashCopy Mappings

Figure 9-4 shows the Copy Services function icon.

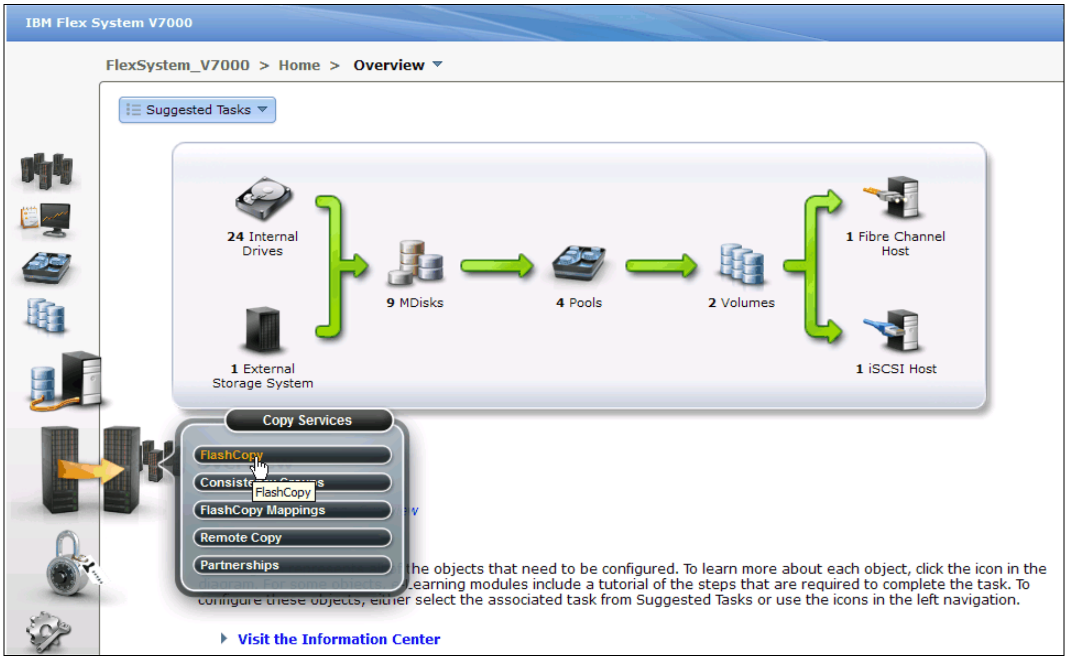


Figure 9-4 Copy Services function icon

Most of the actions to manage the FlashCopy mapping could be done both in the FlashCopy window or the FlashCopy Mappings windows, although the quick path to create FlashCopy presets can only be found in the FlashCopy window.

Click **FlashCopy** in the Copy Services function icon menu, and the FlashCopy window opens (Figure 9-5). In the FlashCopy window, the FlashCopy mappings are organized by volumes.

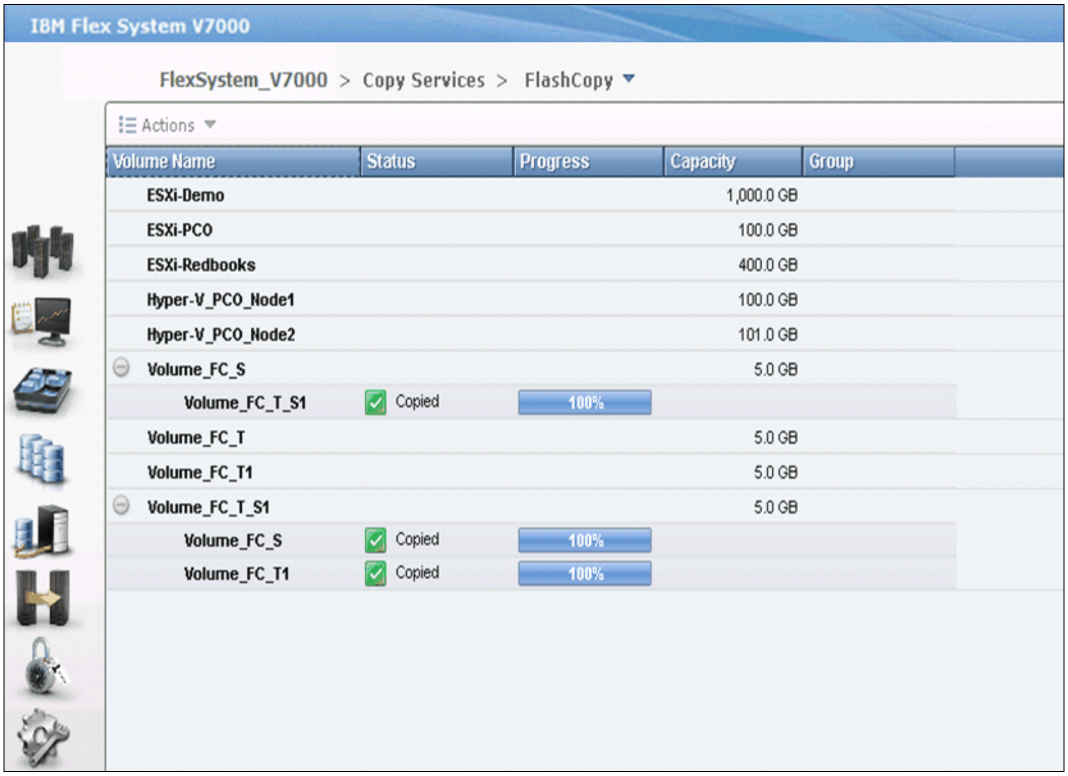


Figure 9-5 FlashCopy window

Click **FlashCopy Mappings** in the Copy Services function icon menu, and the FlashCopy Mappings window opens (Figure 9-6). In the FlashCopy Mappings window, the FlashCopy mappings are listed one by one.

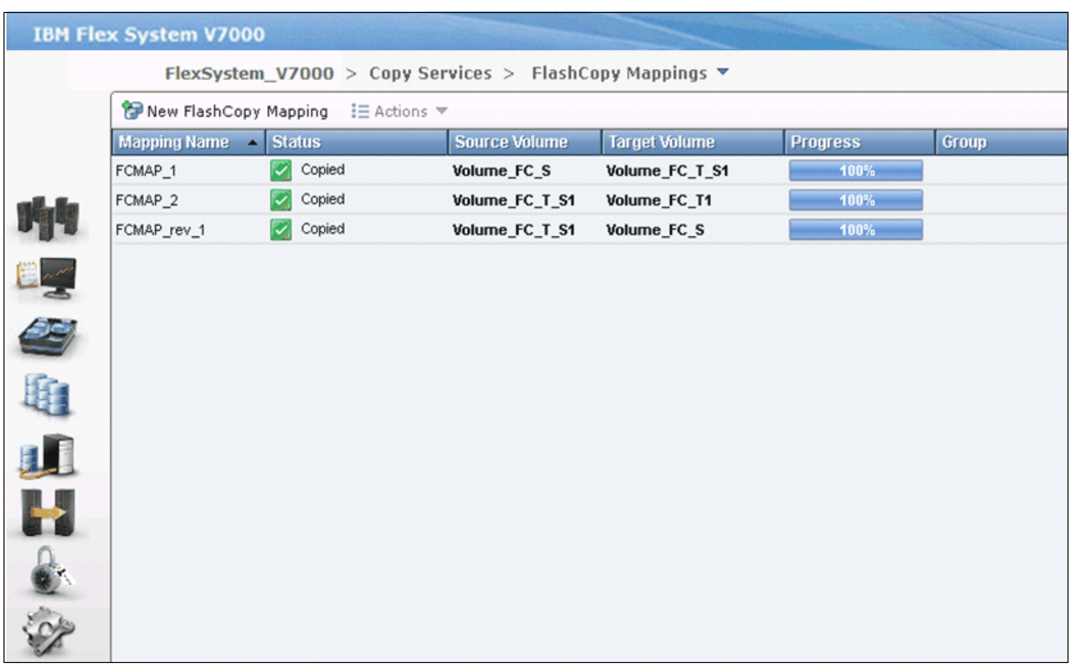


Figure 9-6 FlashCopy Mappings window

The Consistency Groups window could be used to manage the consistency groups for FlashCopy mappings. Click **Consistency Groups** in the Copy Services function icon menu, and the Consistency Groups window opens (Figure 9-7).

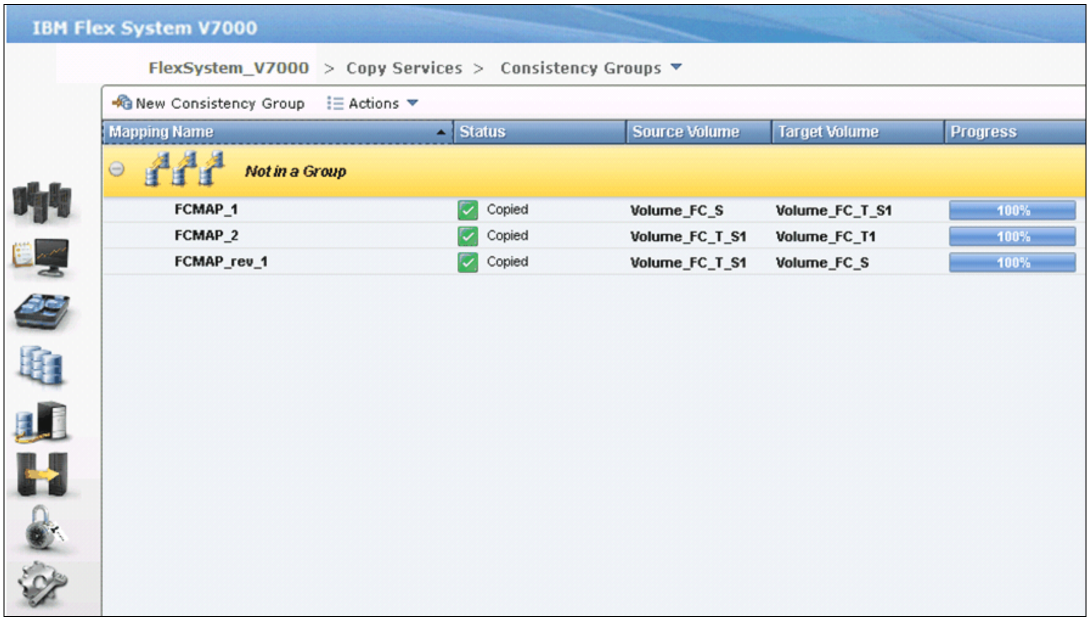


Figure 9-7 Consistency Groups window

Quick path to create FlashCopy presets

It is easy to create FlashCopy using the presets in the FlashCopy window.

Creating a snapshot

In the FlashCopy window, choose a volume and click **New Snapshot** from the **Actions** drop-down menu (Figure 9-8).

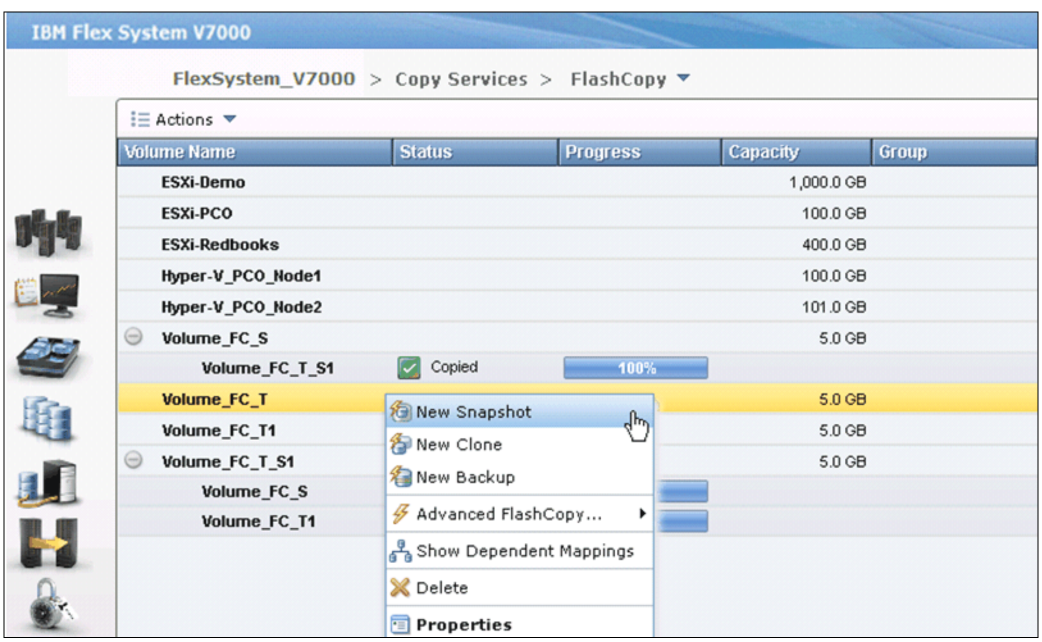


Figure 9-8 Create a snapshot from the preset

You now have a snapshot volume for the volume you choose.

Creating a clone

In the FlashCopy window, choose a volume and click **New Clone** from the **Actions** drop-down menu (Figure 9-9).

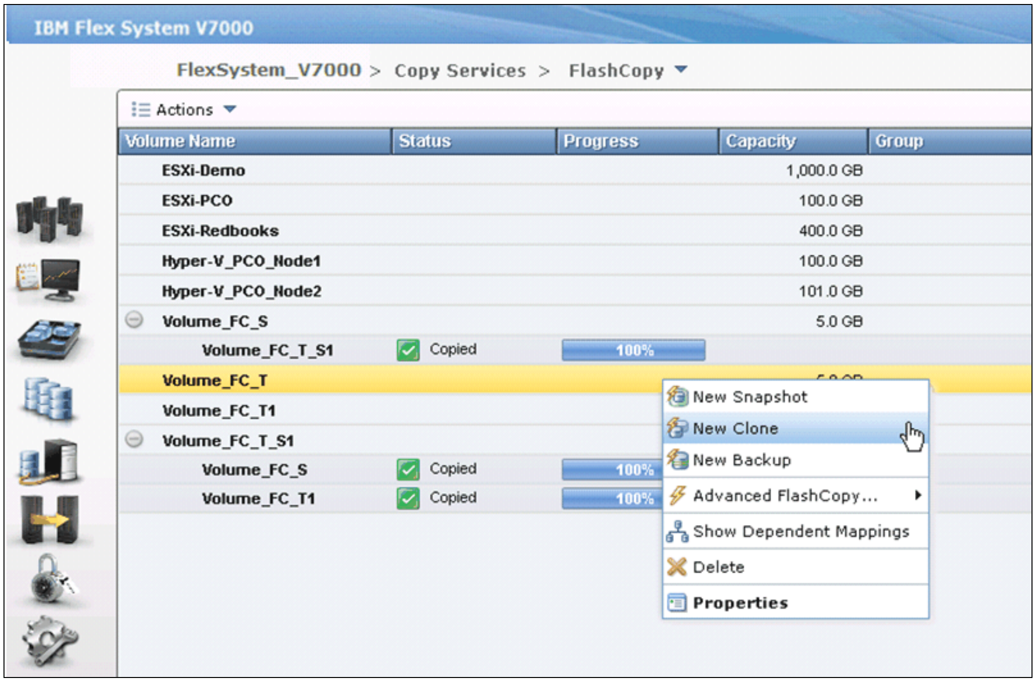


Figure 9-9 Create a clone from the preset

You now have a clone volume for the volume you choose.

Creating a backup

In the FlashCopy window, choose a volume and click **New Backup** from the **Actions** drop-down menu (Figure 9-10).

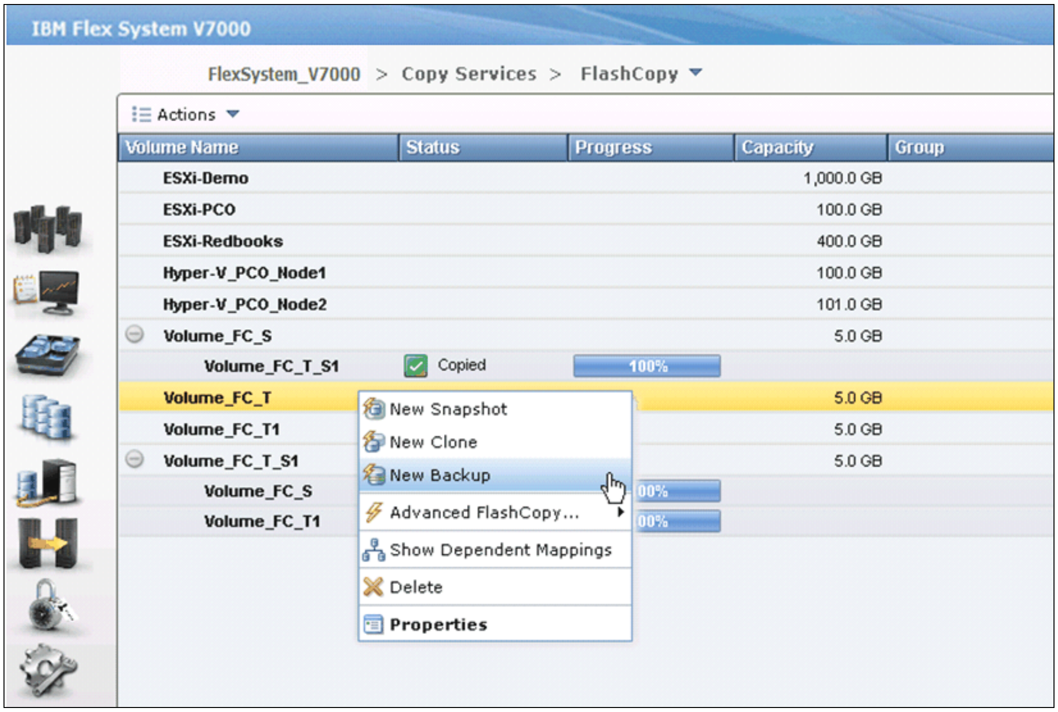


Figure 9-10 Create a backup from the preset

You now have a backup volume for the volume you choose.

Now, in the FlashCopy window, you find three FlashCopy target volumes under the source volume (Figure 9-11). The progress bars behind the target volumes indicate the copy progress in percentage. The copy progress remains 0% for snapshot; there is no change happening at the moment because only the changed data is being copied. At the same time, the copy progresses for clone and backup keep increasing.

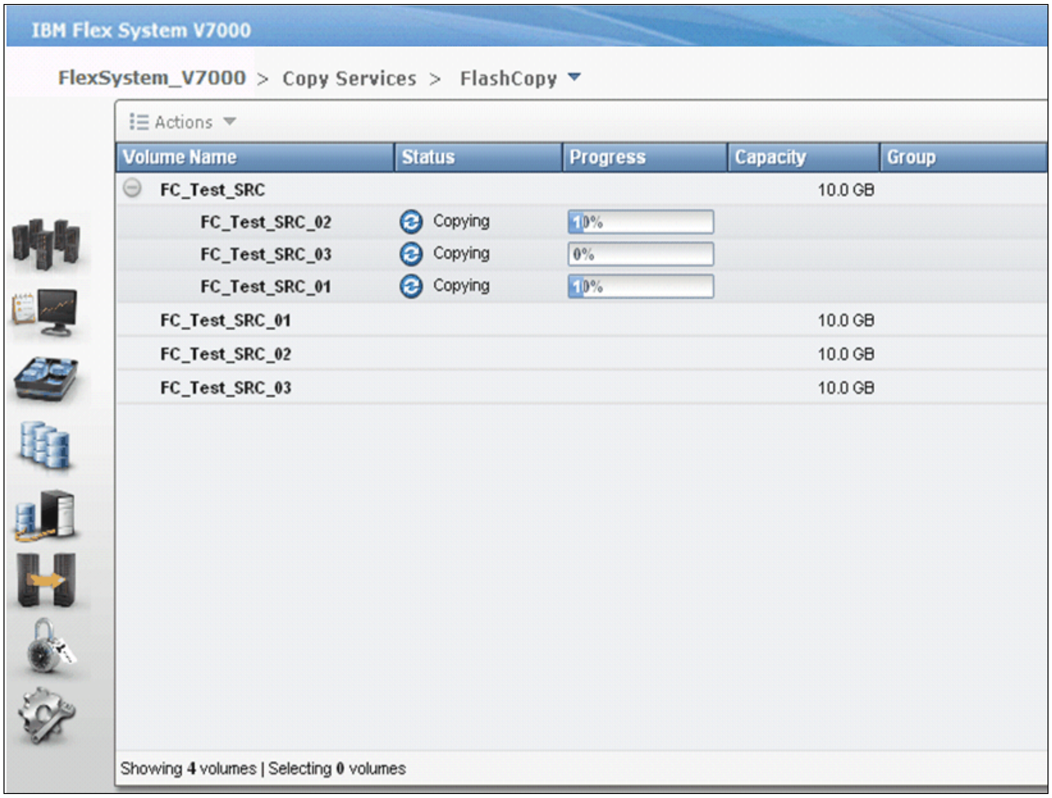


Figure 9-11 FlashCopy in progress viewed in the FlashCopy window

The copy progress can be also found under the Running Tasks status indicator (Figure 9-12). This view is slightly different than the Remote Copy Running Tasks status indicator.

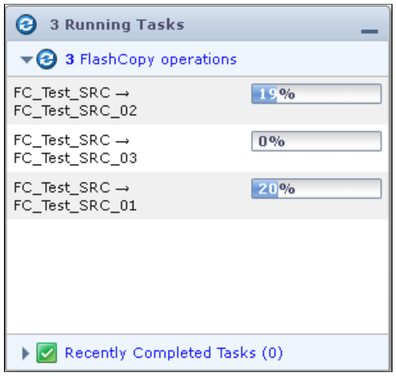


Figure 9-12 The FlashCopy progress

After the copy progresses complete, you find the FlashCopy with the clone preset has been deleted automatically (Figure 9-13).

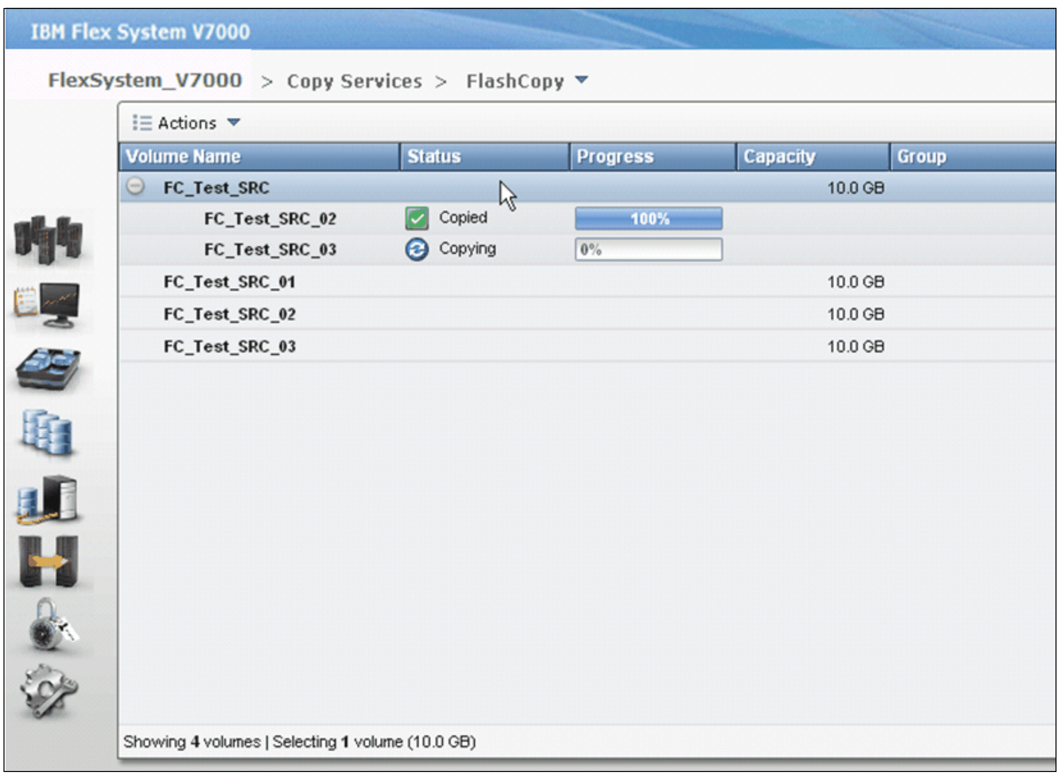


Figure 9-13 FlashCopy progresses complete

9.2.5 Managing FlashCopy mapping

The FlashCopy presets cover the most used FlashCopy configurations for general situations. However, customized FlashCopy mappings are still necessary in some complicated scenarios.

Creating FlashCopy mappings

You can create FlashCopy mappings through the FlashCopy window. Select the volume you want to be the source volume for the FlashCopy mapping and click **Advanced FlashCopy...** from the **Actions** drop-down menu (Figure 9-14).

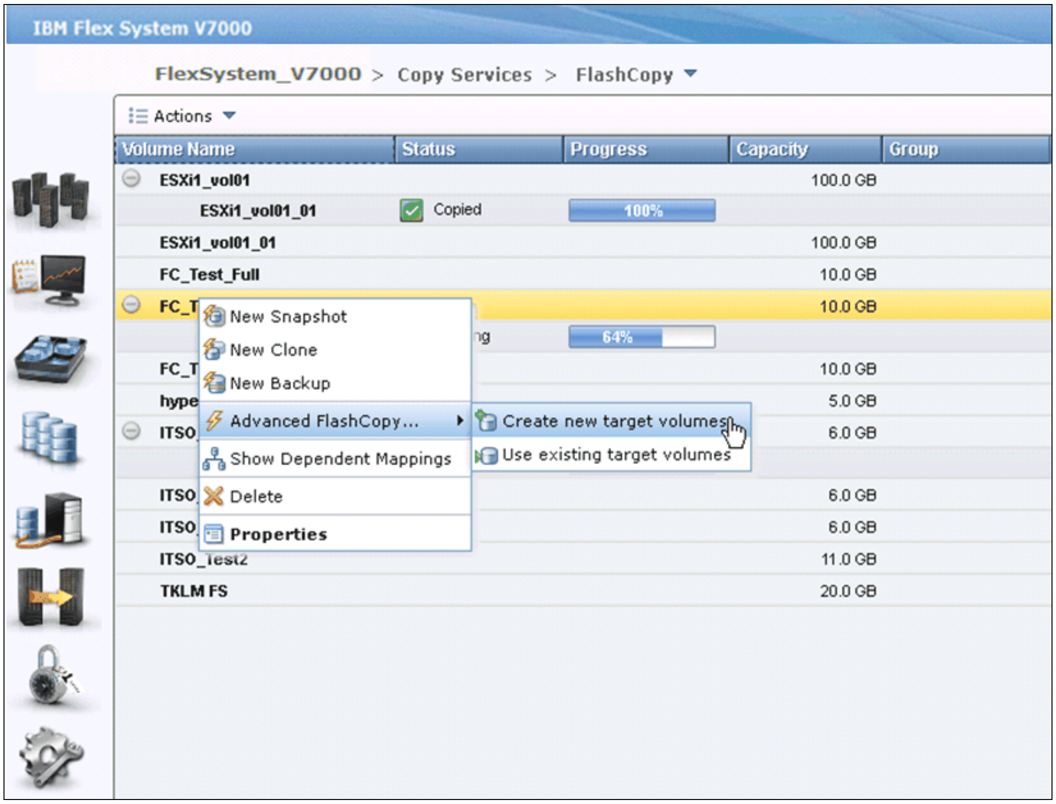


Figure 9-14 Create advanced FlashCopy

You have two options:

- Click **Create new target volumes**, if you have not created the target volume yet.

  a. After you click **Create new target volumes**, the wizard lets you choose the preset, but no matter which preset you choose, you can modify the setting of the FlashCopy mapping. Choose one preset that has the most similar configuration to the one required, and click **Advanced Settings** to make any appropriate adjustments on the properties (Figure 9-15). Figure 9-16 on page 361 shows the default setting.

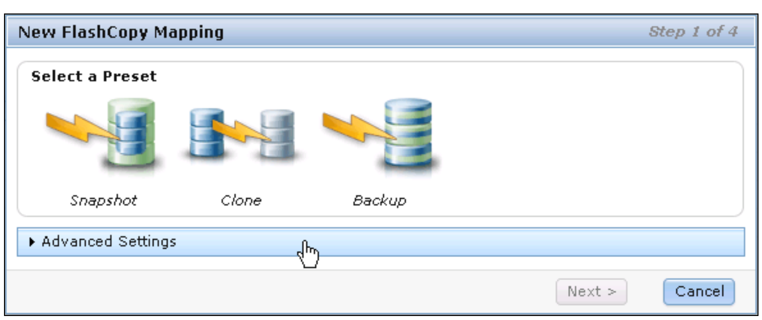


Figure 9-15 Choose a preset most similar to your requirement

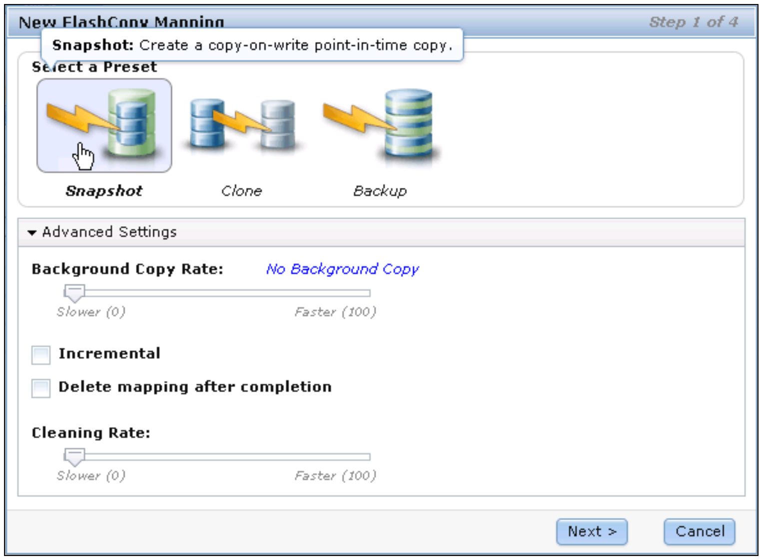


Figure 9-16 Default setting for the snapshot preset

For example, if the snapshot preset has been chosen, the default settings can be found when you click **Advanced Settings**, which are:

- Background Copy: 0
- Incremental: No
- Auto Delete after completion: No
- Cleaning Rate: 0

It is the same with clone preset, and the default settings of clone preset can be found when you click **Advanced Settings**, after you select clone preset (Figure 9-17), which includes:

- Background Copy: 50
- Incremental: No
- Auto Delete after completion: Yes
- Cleaning Rate: 50

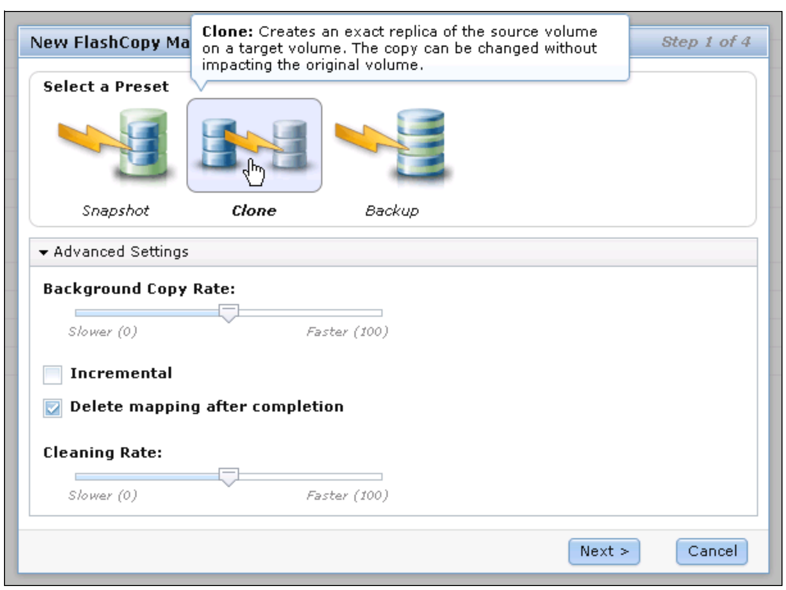


Figure 9-17 Default settings for the clone preset

It is the same with backup preset, and the default settings of backup preset are found when you click **Advanced Settings**, after you select the backup preset (Figure 9-18), which includes:

- Background Copy: 50
- Incremental: Yes
- Auto Delete after completion: No
- Cleaning Rate: 50

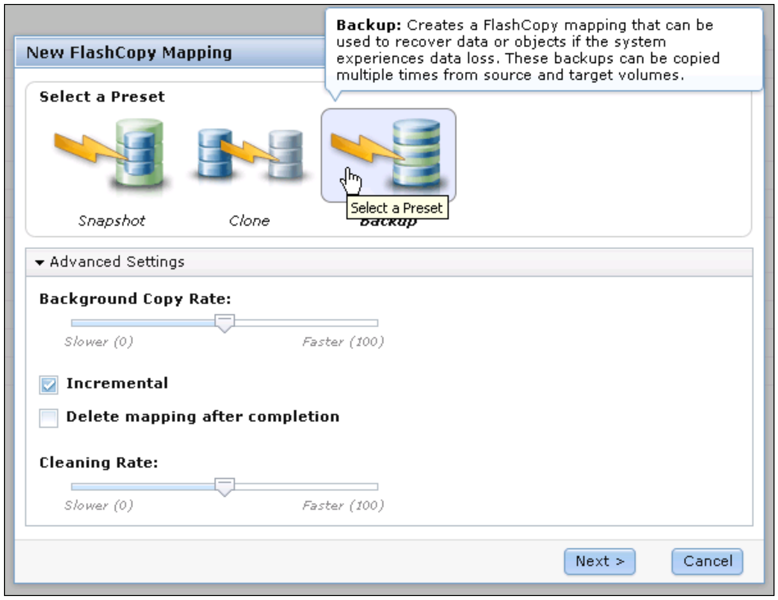


Figure 9-18 Default settings for the backup preset

Change the settings of the FlashCopy mapping according to your requirements and click **Next**.

b. In the next step, you could add your FlashCopy mapping to a consistency group (Figure 9-19). If the consistency group is not ready, the FlashCopy mapping can be added to the consistency group afterward. Click **Next** to continue.

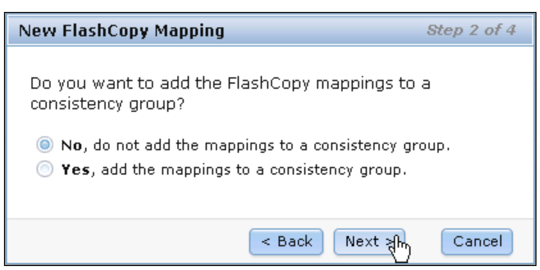


Figure 9-19 Add FlashCopy mapping to a consistency group

c. You may choose from which storage pool you want to create your volume. As shown in Figure 9-20 on page 363, you can select the same storage pool that is used by the source volume. Click **Next** to continue.

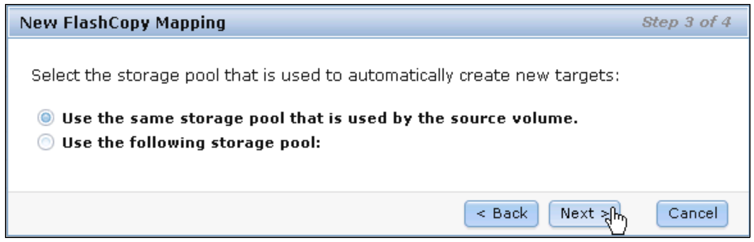


Figure 9-20 Choose use the same storage pool with the source volume

d. You can also specify some other storage pool for your new volume (Figure 9-21). Click **Next** to continue.

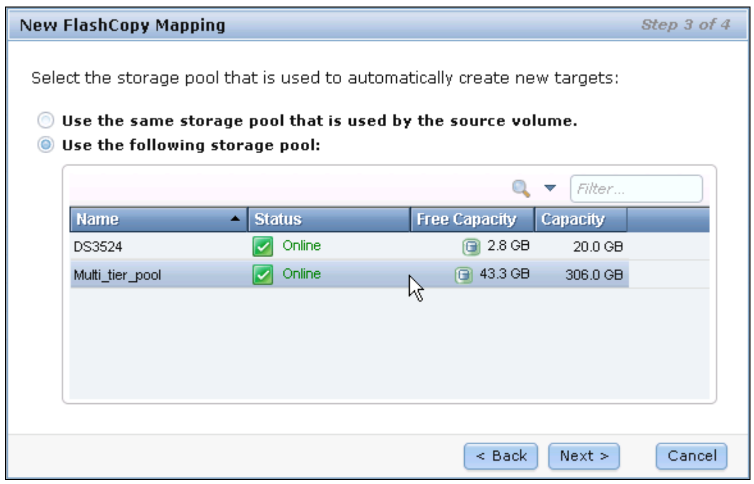


Figure 9-21 Choose another storage pool to create the volume

e. You are prompted to select how you want to manage volume capacity. **Create a generic volume** is your default choice. If you select a thin-provisioned volume, you get additional options (Figure 9-22).

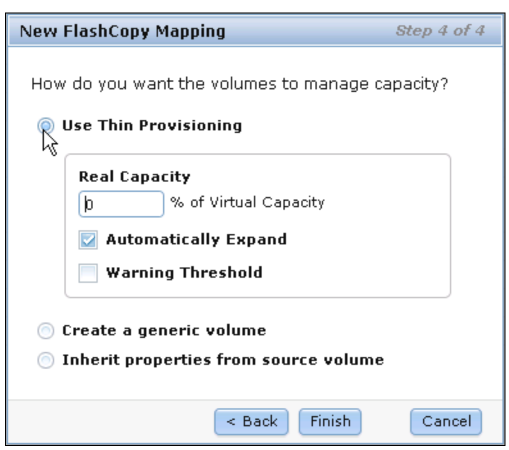


Figure 9-22 Create a thin provisioned target volume

If you chose the clone or backup preset at the beginning of this wizard, **Create a generic volume** is your default choice (Figure 9-23 on page 364).

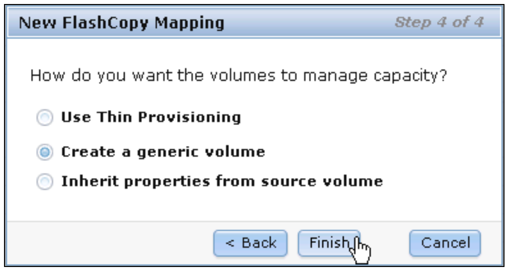


Figure 9-23 The default choice when creating a FlashCopy mapping

f. Click **Finish** when you make your decision, and the FlashCopy mapping is created on your volume with a new target (Figure 9-24). The status of the newly created FlashCopy mapping is Idle; it could be started, as described in “Starting a FlashCopy mapping” on page 367.

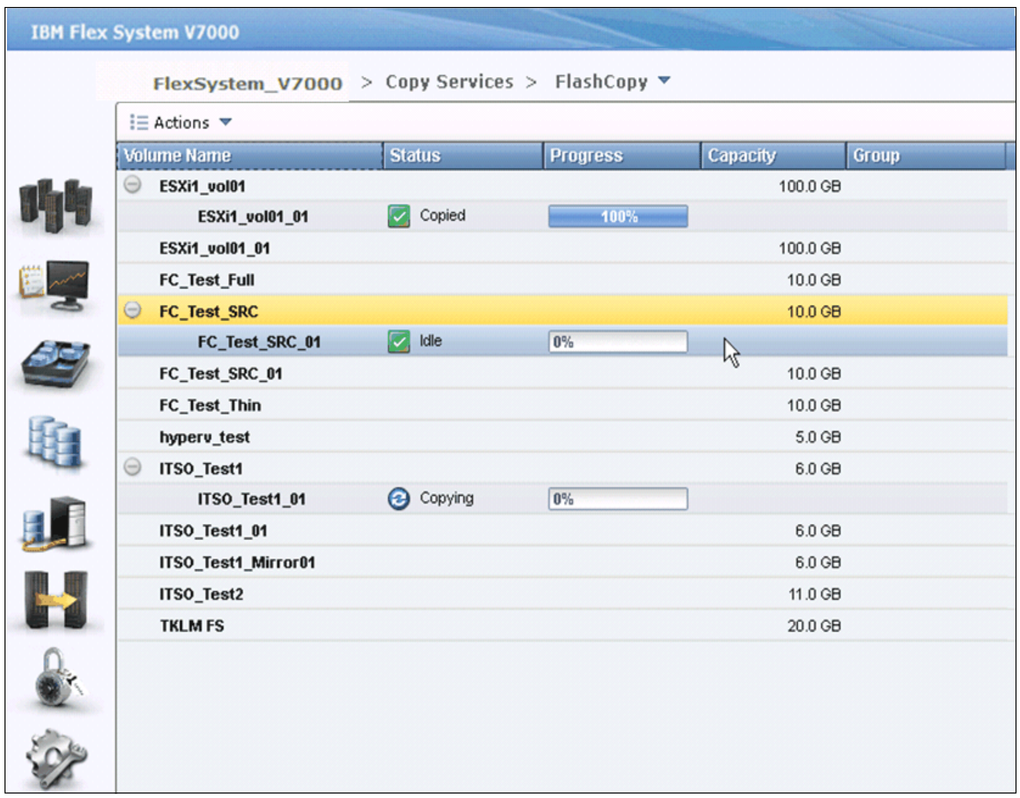


Figure 9-24 New FlashCopy mapping has been created with a new target

- In the Advanced FlashCopy menu, if you already have candidate target volumes, select **Use existing target volumes** (Figure 9-25).

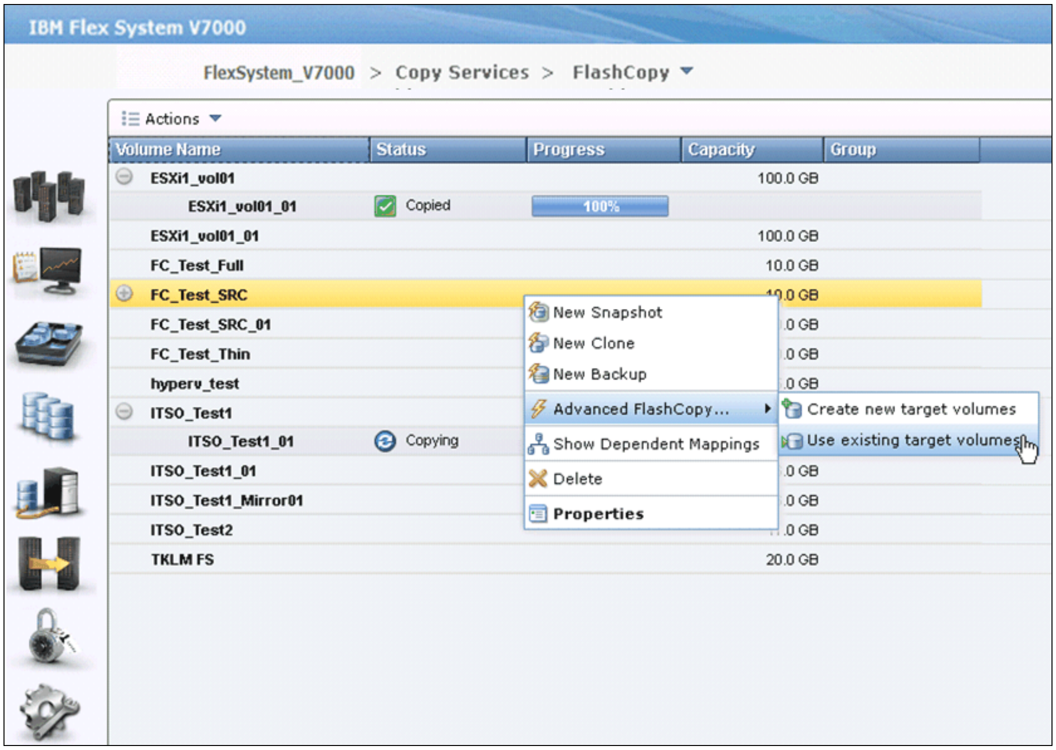


Figure 9-25 Create FlashCopy mapping using existing target volume

a. You need to choose the target volume for the source volume you selected. Select the target volume in the drop-down menu in the right pane of the window and click **Add** (Figure 9-26).

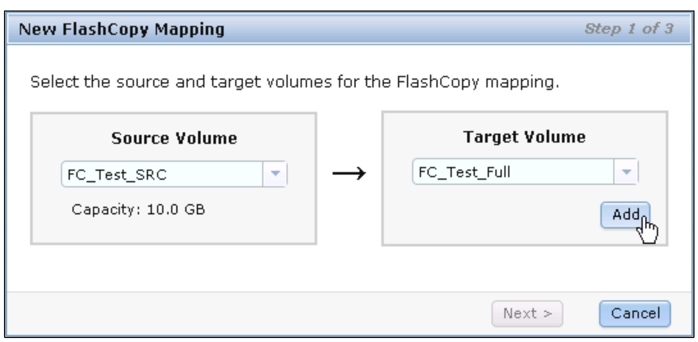


Figure 9-26 Select the target volume

b. After you click **Add**, the FlashCopy mapping is listed (Figure 9-27). Click the red X if the FlashCopy mapping is not the one you want to create. If the FlashCopy mapping is what you want, click **Next** to continue.

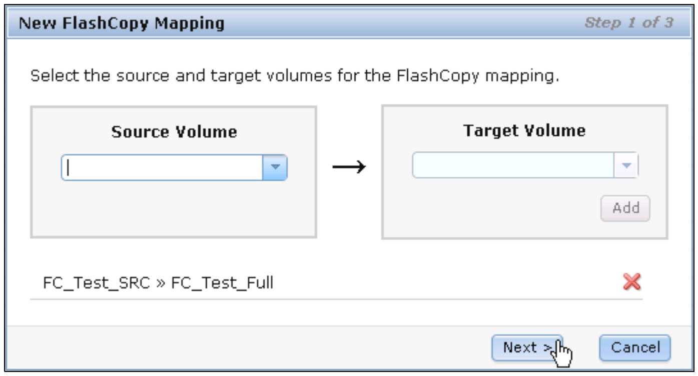


Figure 9-27 Add FlashCopy mapping

c. Select the preset and adjust the settings (Figure 9-28). Make sure that the settings meet your requirements, and click **Next**.

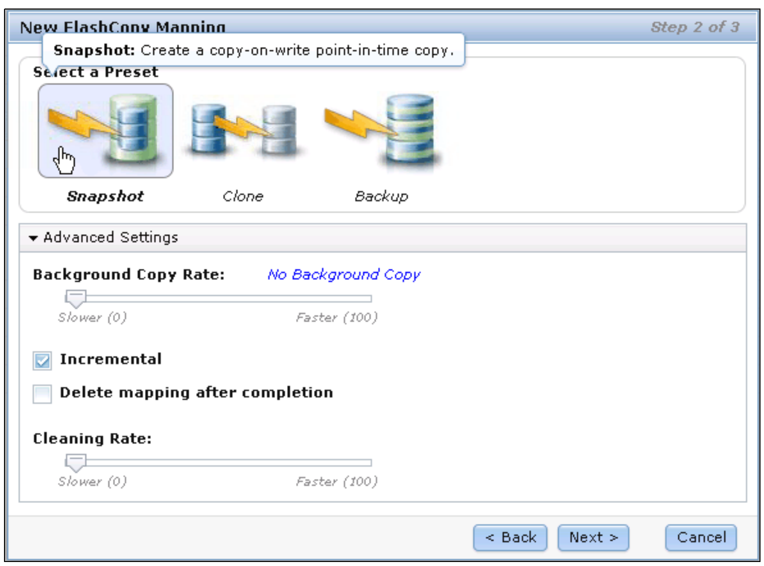


Figure 9-28 Select a preset and make your adjustments

d. Now you can add the FlashCopy mapping to a consistency group (Figure 9-29). Click **Finish** and the FlashCopy mapping is created with the status of Idle.

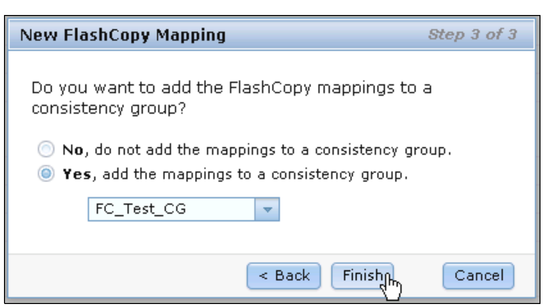


Figure 9-29 Select a consistency group to add the FlashCopy mapping to

You can also create the FlashCopy mappings in the FlashCopy Mapping window by clicking **New FlashCopy Mapping** at the upper left, as shown in Figure 9-30.

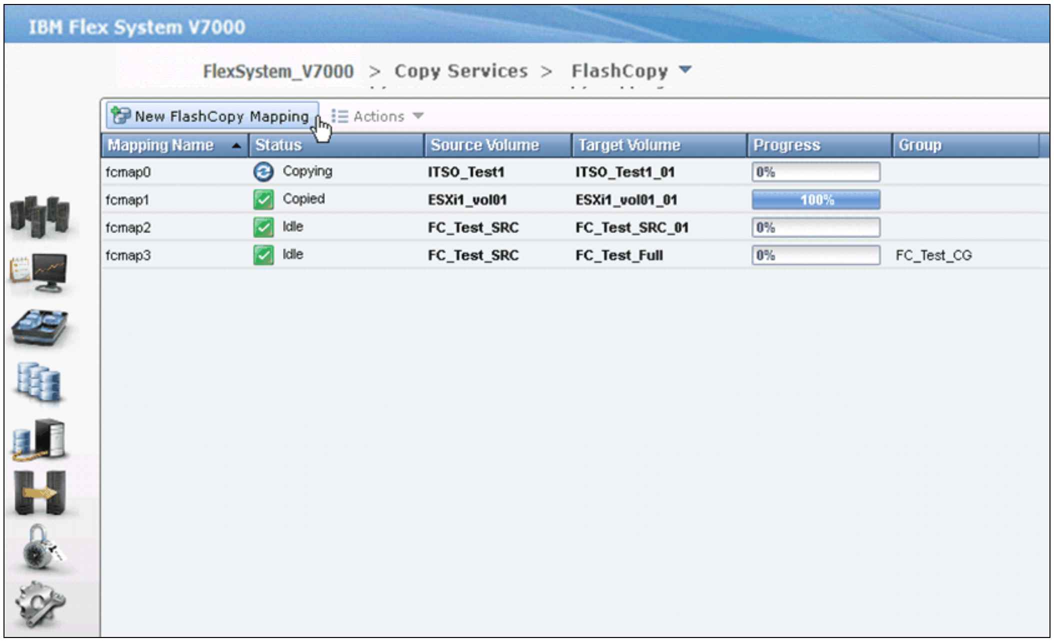


Figure 9-30 Create a FlashCopy mapping in the FlashCopy Mappings window

A wizard opens to guide you through the creation of a FlashCopy mapping, and the steps are the same as creating an advanced FlashCopy mapping using existing target volumes in the FlashCopy window.

Starting a FlashCopy mapping

Most of the actions of FlashCopy mapping can be taken both in the FlashCopy window or the FlashCopy Mapping window. For the actions available in both windows, we show, in the following sections, the steps in the FlashCopy window, although the steps would be the same if you were to use the FlashCopy Mapping window.

You can start the mapping by selecting the FlashCopy target volume in the FlashCopy window and selecting the **Start** option from the **Actions** drop-down menu (Figure 9-31 on page 368). The status of the FlashCopy mapping changes from Idle to Copying.

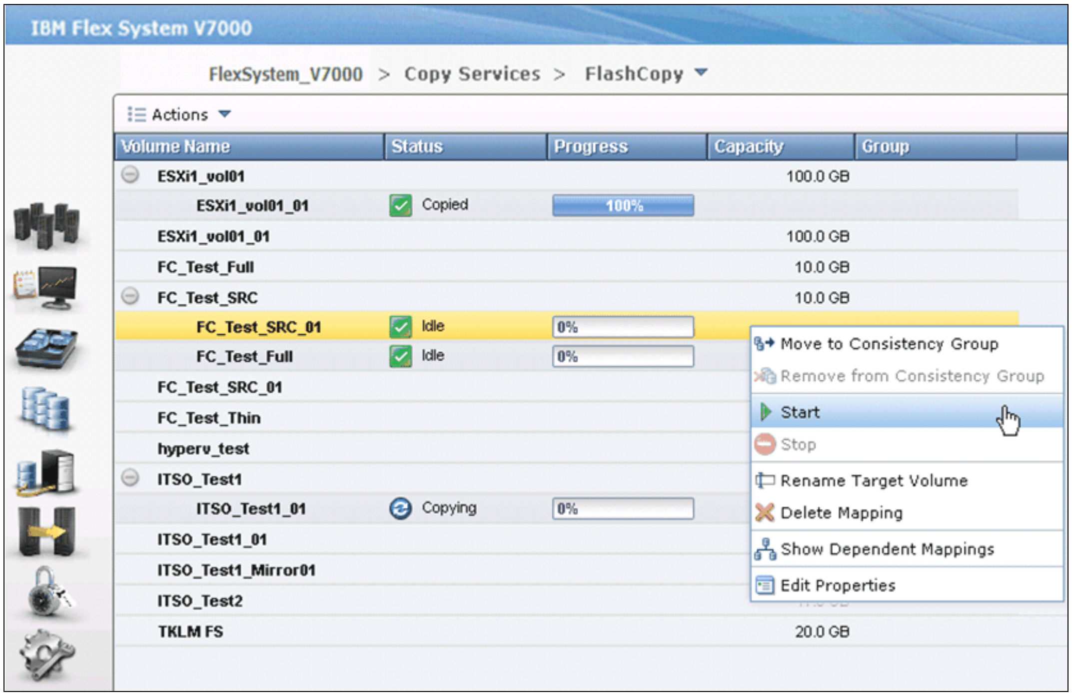


Figure 9-31 Start FlashCopy mapping

Stopping a FlashCopy mapping

The FlashCopy mapping can be stopped by selecting the FlashCopy target volume in the FlashCopy window and clicking the **Stop** option from the **Actions** drop-down menu (Figure 9-32). After the stopping process completes, the status of the FlashCopy mapping is changed to Stopped.

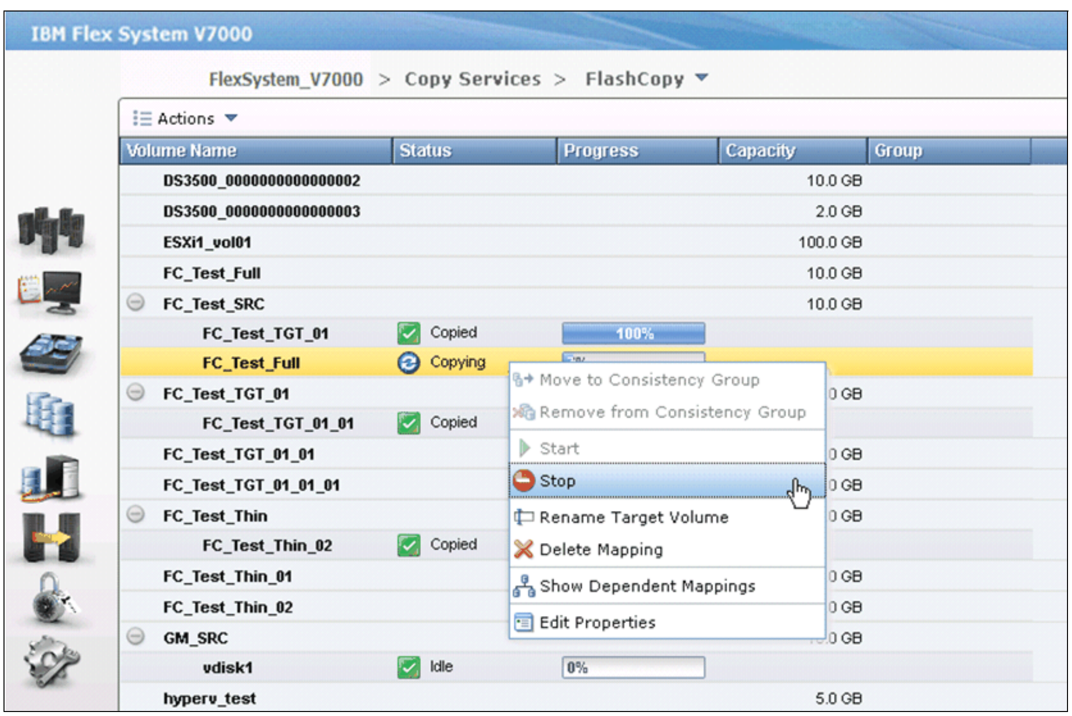


Figure 9-32 Stopping a FlashCopy mapping

Renaming the target volume

If the FlashCopy target volumes were created automatically by IBM Flex System V7000 Storage Node system, the name of the target volumes is the source volume name plus a suffix containing numbers. The name of the target volumes could be changed to be more meaningful in your environment.

To change the name of the target volume, select the FlashCopy target volume in the FlashCopy window and click the **Rename Target Volume** option from the **Actions** drop-down menu (Figure 9-33).

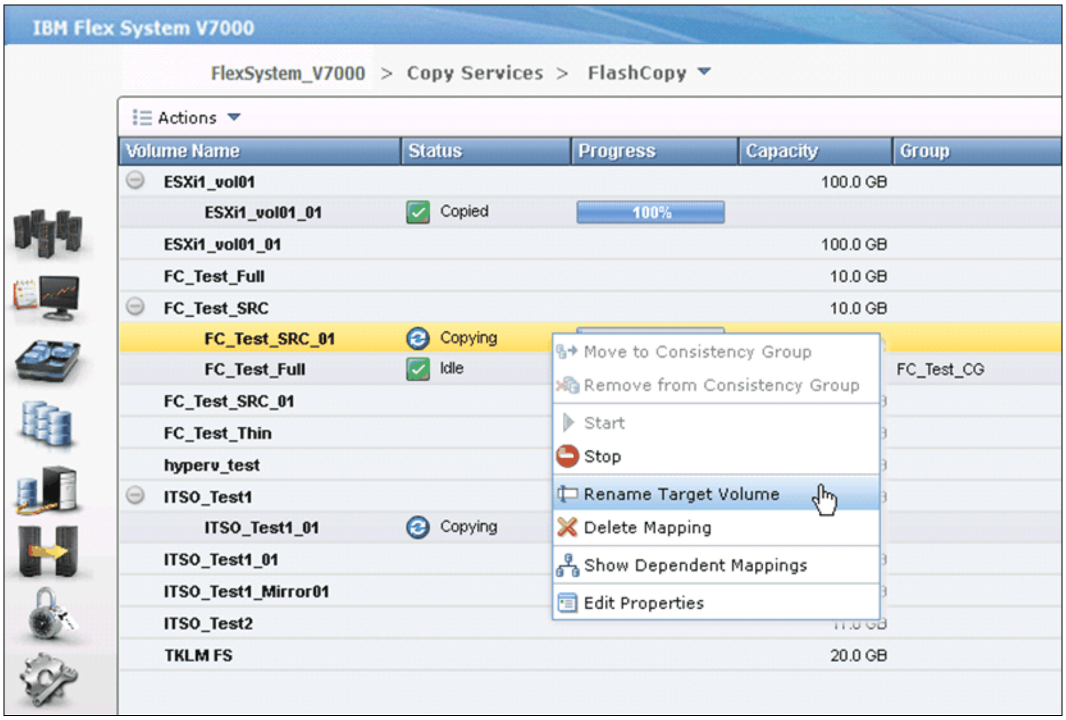


Figure 9-33 Rename a target volume

Input your new name for the target volume (Figure 9-34). Click **Rename** to finish.

Figure 9-34 Rename a target volume

Renaming a FlashCopy mapping

The FlashCopy mappings are created with names that begin with fcmap. The name of FlashCopy mappings can be changed to be more meaningful to you.

To change the name of a FlashCopy mapping, select the FlashCopy mapping in the FlashCopy Mappings window and click the **Rename Mapping** option from the **Actions** drop-down menu (Figure 9-35).

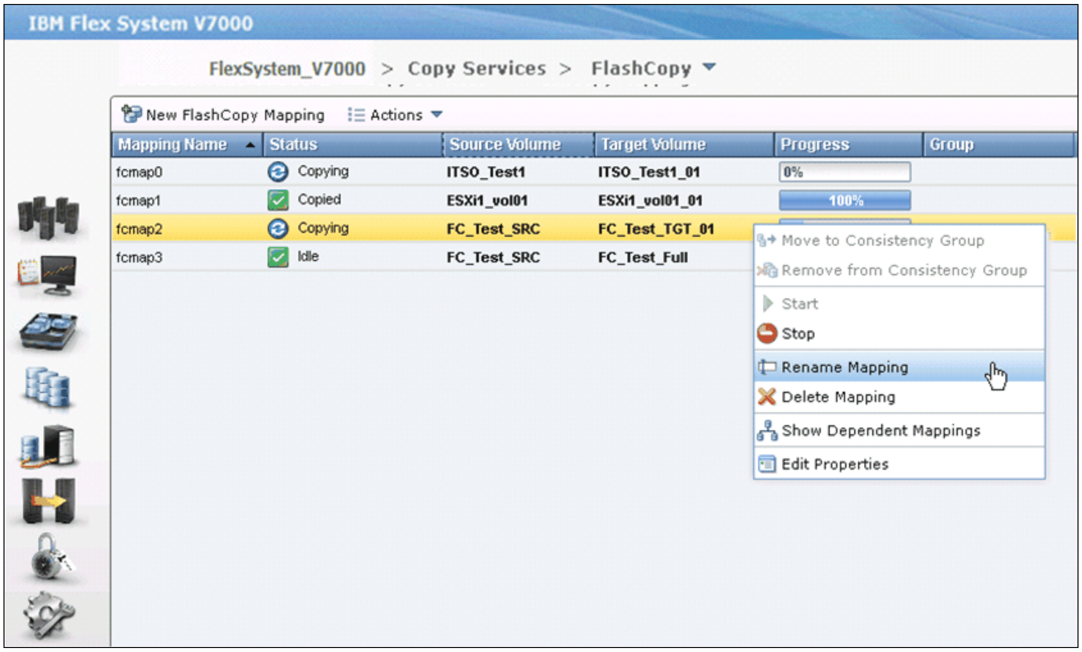


Figure 9-35 Rename a FlashCopy mapping

You need to enter your new name for the FlashCopy mapping (Figure 9-36). Click **Rename** to finish.

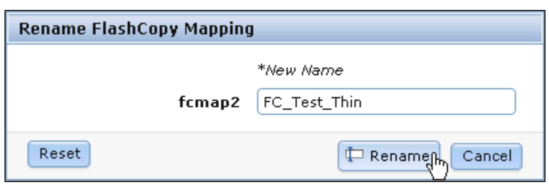


Figure 9-36 Enter a new name for the FlashCopy mapping

Deleting a FlashCopy mapping

The FlashCopy mapping can be deleted by selecting the FlashCopy target volume in the FlashCopy window and clicking the **Delete Mapping** option from the **Actions** drop-down menu (Figure 9-37).

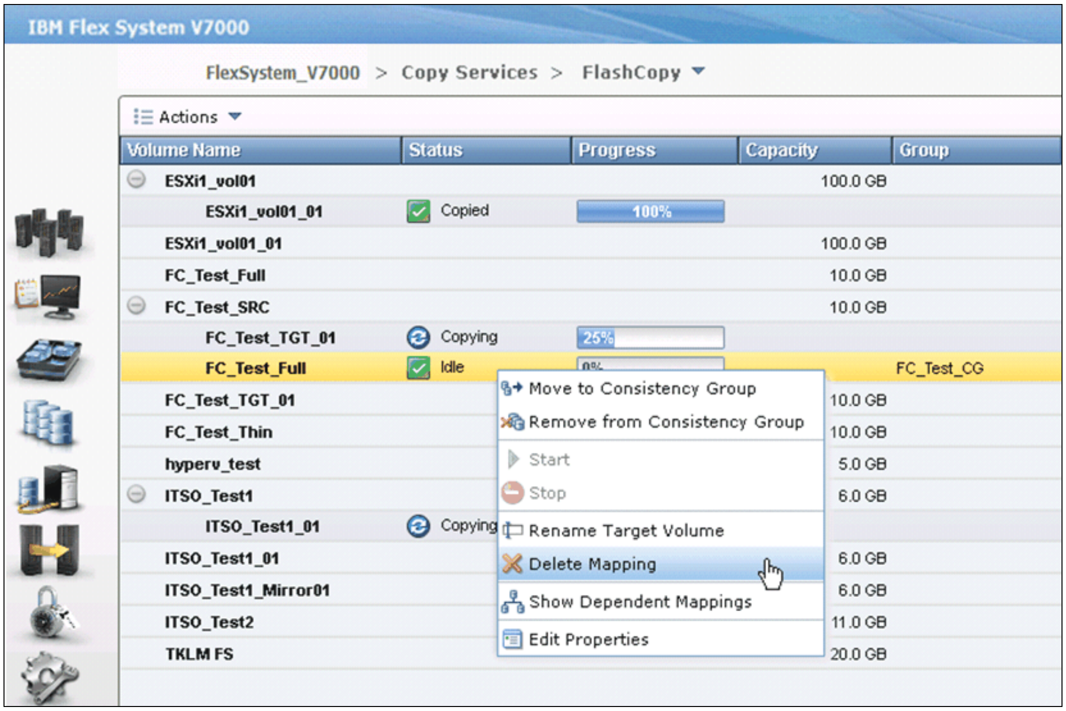


Figure 9-37 Select Delete Mapping

FlashCopy Mapping state: If the FlashCopy mapping is in the Copying state, it must be stopped before being deleted.

You need to confirm your action to delete FlashCopy mappings in the window that opens (Figure 9-38). Verify the number of FlashCopy mappings that you need to delete, and if you want to delete the FlashCopy mappings while the data on the target volume is inconsistent with the source volume, select the check box. Click **Delete** and your FlashCopy mapping is removed.

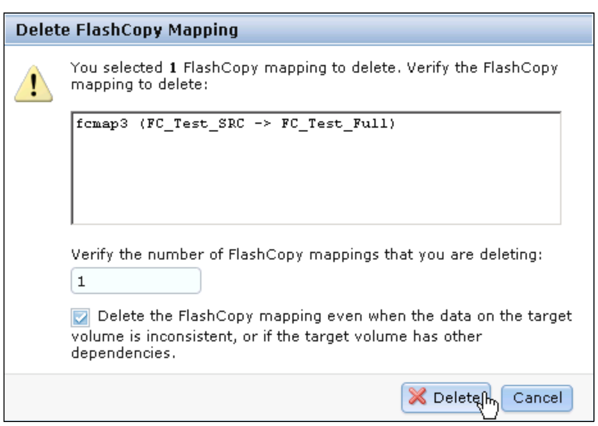


Figure 9-38 Confirm the deletion of FlashCopy mappings

Deleting FlashCopy mapping: Deleting the FlashCopy mapping does not delete the target volume. If you need to reclaim the storage space occupied by the target volume, you need to delete the target volume manually.

Showing dependent mapping

You can show the FlashCopy mapping dependency by selecting the FlashCopy source volume in the FlashCopy window and clicking the **Show Dependent Mappings** option from the **Actions** drop-down menu (Figure 9-39).

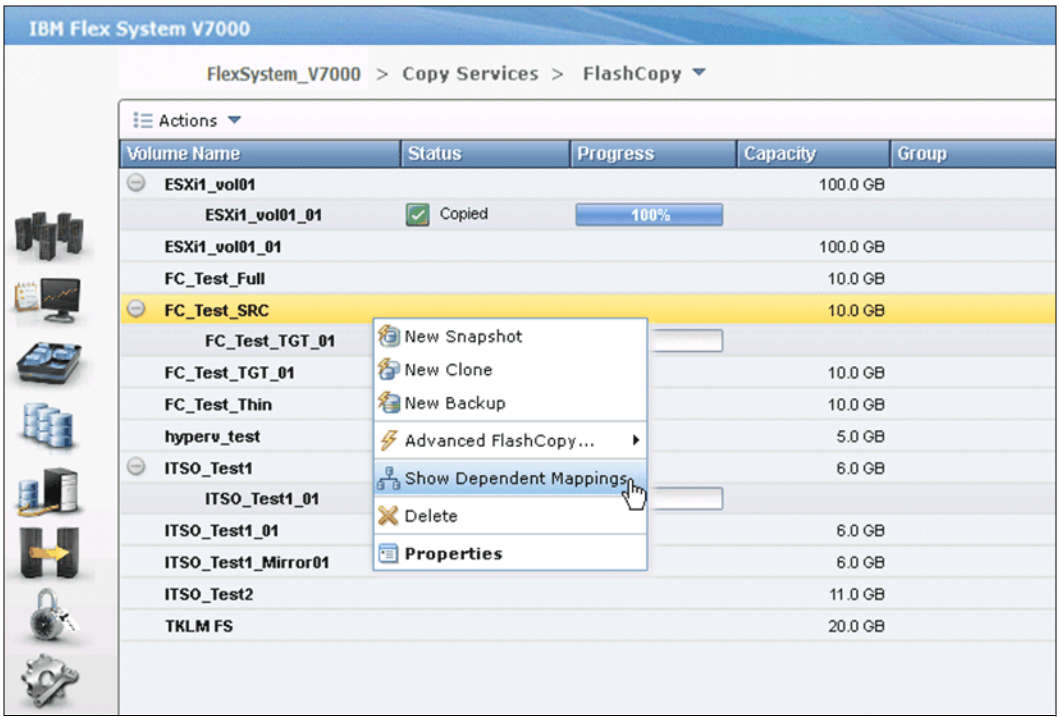


Figure 9-39 Show Dependent Mappings menu

The FlashCopy mapping dependency tree opens (Figure 9-40).

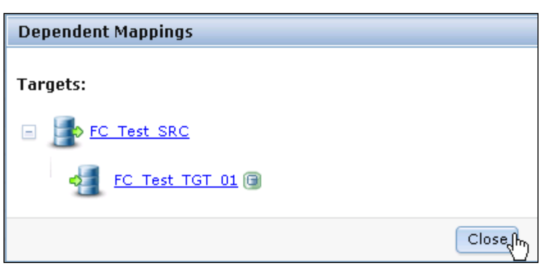


Figure 9-40 FlashCopy mapping dependency

Editing properties

The background copy rate and cleaning rate can be changed after the FlashCopy mapping has been created by selecting the FlashCopy target volume in the FlashCopy window and clicking the **Edit Properties** option from the **Actions** drop-down menu (Figure 9-41).

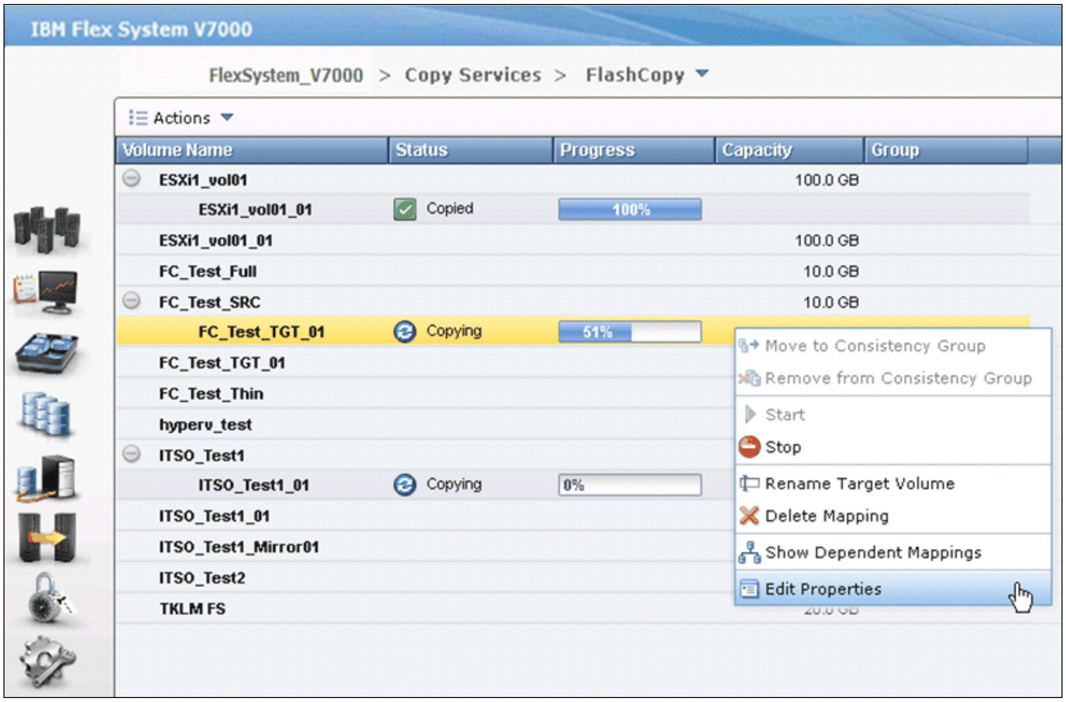


Figure 9-41 Edit Properties menu

You can then modify the value of the background copy rate and cleaning rate by moving the pointers on the bars (Figure 9-42). Click **Save** to save changes.

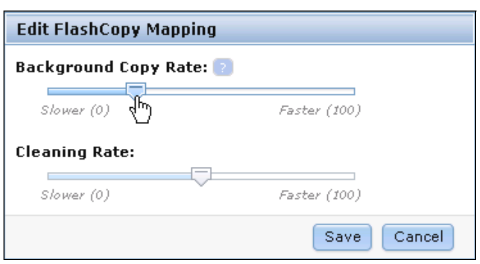


Figure 9-42 Change the copy rate

Managing a FlashCopy consistency group

FlashCopy consistency groups can be managed by clicking the **Consistency Groups** menu under the Copy Services function icon (Figure 9-43 on page 374).

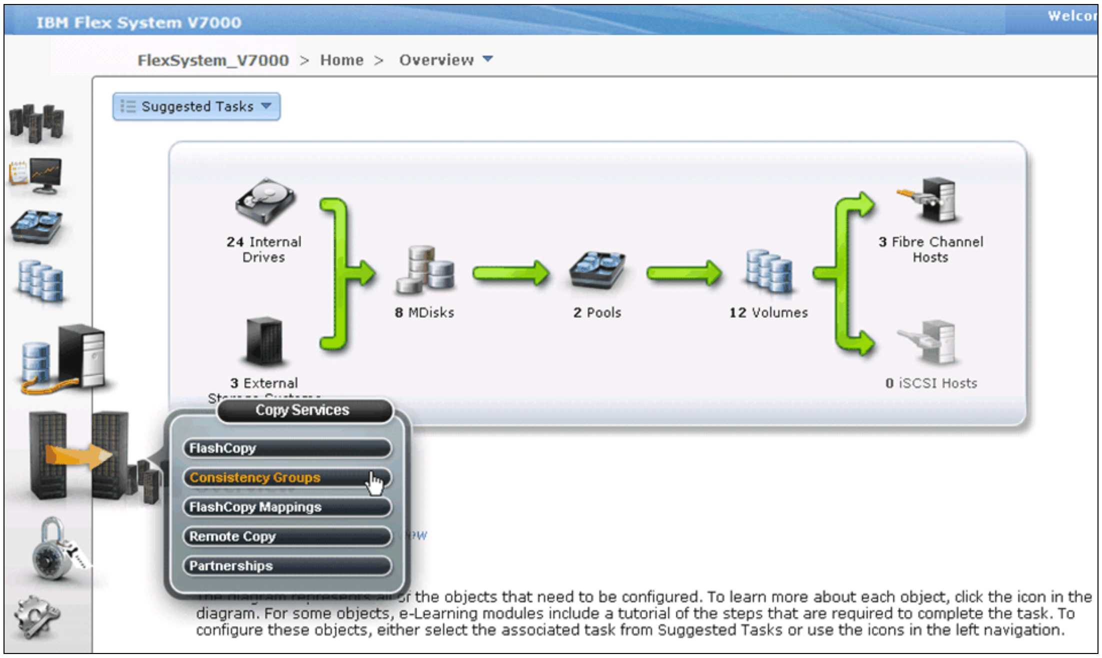


Figure 9-43 Access to the Consistency Groups window

The Consistency Groups window (Figure 9-44) is where you can manage both consistency groups and FlashCopy mappings.

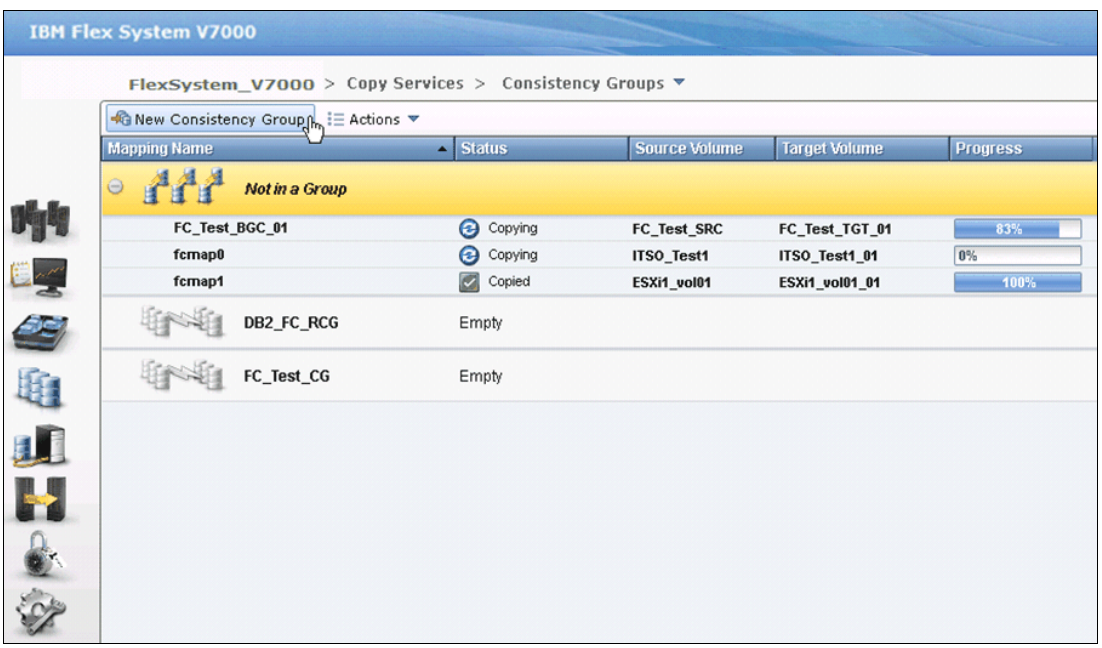


Figure 9-44 Consistency Groups window

In the left pane of the Consistency Groups window, you can list the consistency groups you need. Click **Not in a Group**, and then expand your selection by clicking the plus icon next to **Not in a Group**. All the FlashCopy mappings that are not in any consistency groups are displayed in the lower pane.

In the lower pane of the Consistency Groups window, you can discover the properties of a consistency group and the FlashCopy mappings in it. You can also take actions on any

consistency groups and FlashCopy mappings within the Consistency Groups window, as allowed by their state. All the actions allowed for the FlashCopy mapping are described in 9.2.5, “Managing FlashCopy mapping” on page 359.

Creating a FlashCopy consistency group

To create a FlashCopy consistency group, click **New Consistency Group** at the top of the Consistency Groups window (Figure 9-44 on page 374).

You are prompted to enter the name of the new consistency group (Figure 9-45). Following your naming conventions, enter the name of the new consistency group in the box and click **Create**.

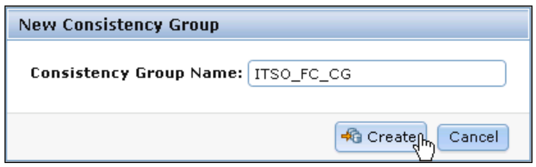


Figure 9-45 Enter the name for the consistency group

After the creation process completes, you find a new consistency group in the left pane of the Consistency Groups window. Select the new consistency group, and you see more detailed information about this consistency group in the right pane (Figure 9-46).

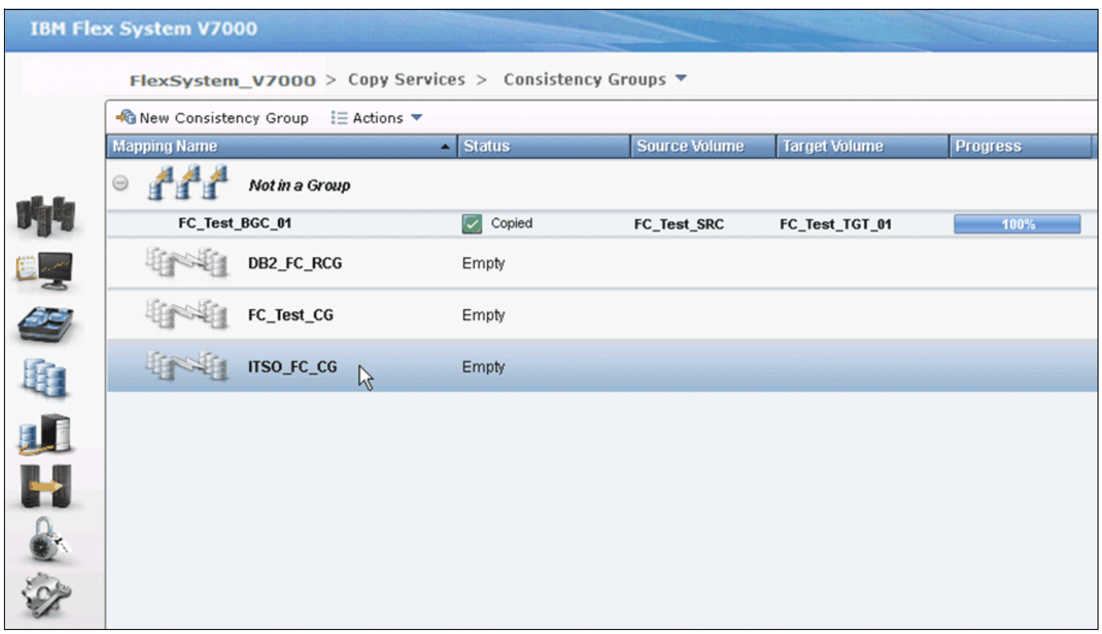


Figure 9-46 New consistency group

You can rename the consistency group by clicking the name of the consistency group in the main pane and enter a name for it (following your naming convention). Next to the name of the consistency group, the state shows that it is now an empty consistency group with no FlashCopy mapping in it.

Adding FlashCopy mappings to a consistency group

Click **Not in a Group** to list all the FlashCopy mappings with no consistency group. You can add FlashCopy mappings to a consistency group by selecting them and clicking the **Move to Consistency Group** option from the **Actions** drop-down menu (Figure 9-47 on page 376).

Selections of a range are performed by highlighting a mapping, pressing and holding the Shift key, and clicking the last item in the range. Multiple selections can be made by pressing and holding the Ctrl key and clicking each mapping individually.

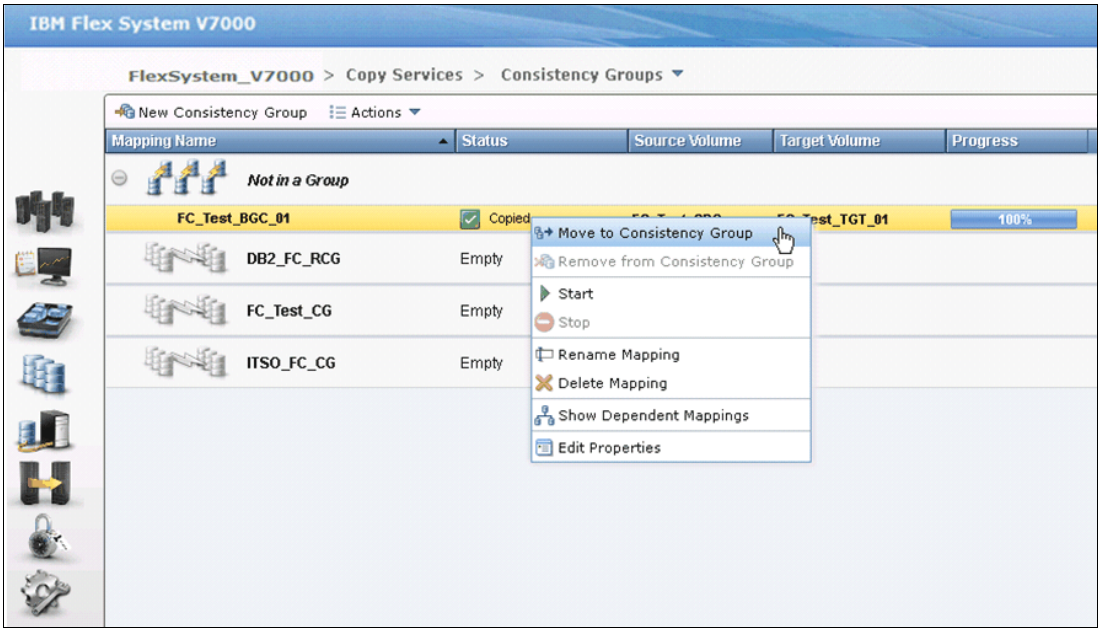


Figure 9-47 Select the FlashCopy mappings to add to a consistency group

You are prompted to specify which consistency group you want to move the FlashCopy mapping into (Figure 9-48). Click **Move to Consistency Group** to continue.

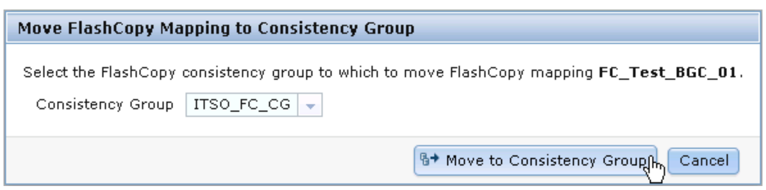


Figure 9-48 Select consistency group

After the action completes, you find that the FlashCopy mappings you selected have been removed from the Not In a Group list of the consistency group you chose.

Starting a consistency group

To start a consistency group, click **Start** from the **Actions** drop-down menu (Figure 9-49).

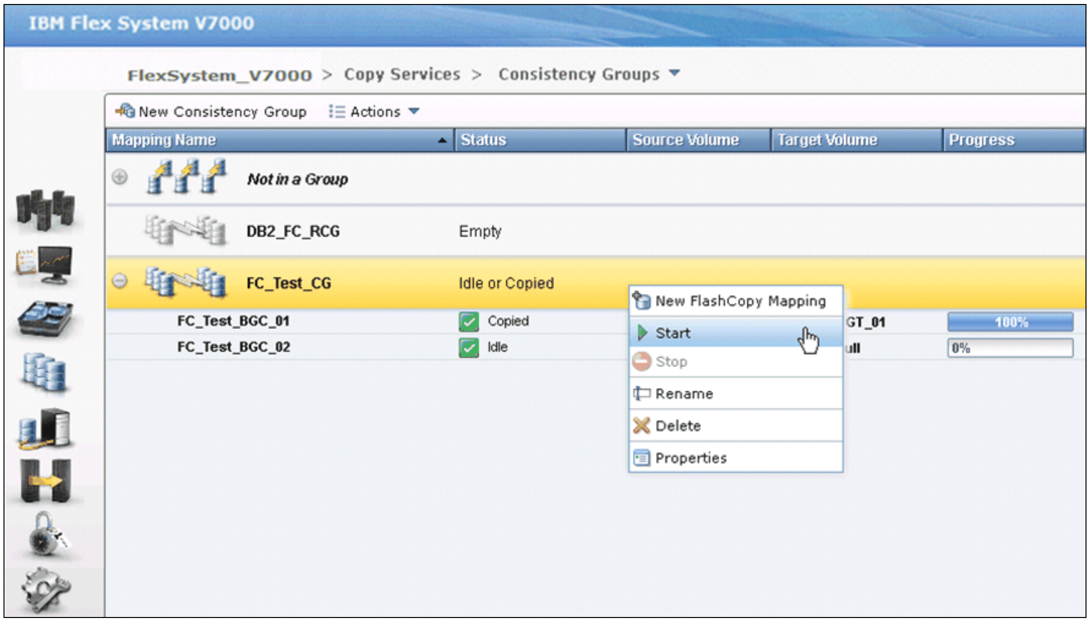


Figure 9-49 Start a consistency group

After you start the consistency group, all the FlashCopy mappings in the consistency group are started at the same time. The state of the FlashCopy changes to Copying (Figure 9-50).

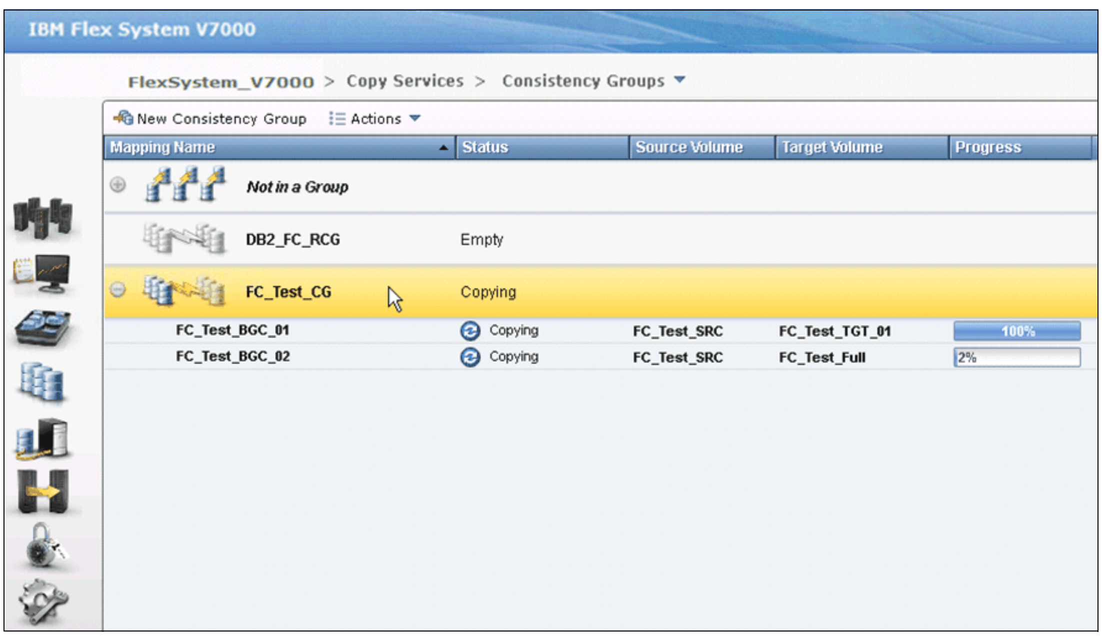


Figure 9-50 Consistency group startup completes

Stopping a consistency group

The consistency group can be stopped by selecting **Stop** from the **Actions** drop-down window (Figure 9-51 on page 378).

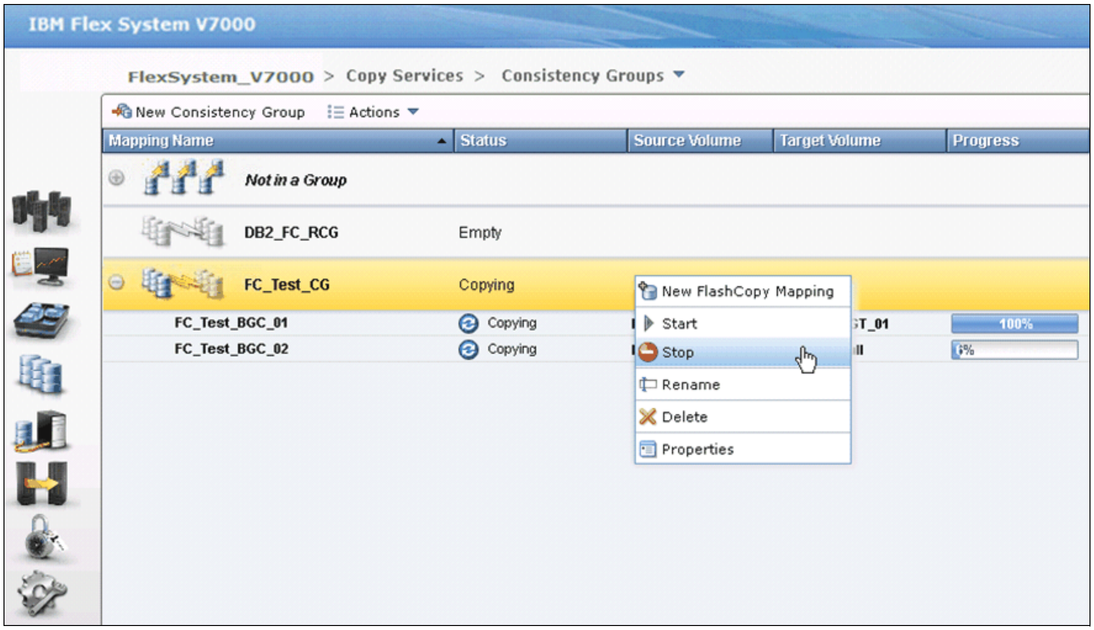


Figure 9-51 Stop a consistency group

After the stop process completes, the FlashCopy mappings in the consistency group are in the Stopped state, and a red “X” icon appears on the function icon of this consistency group to indicate an alert (Figure 9-52).

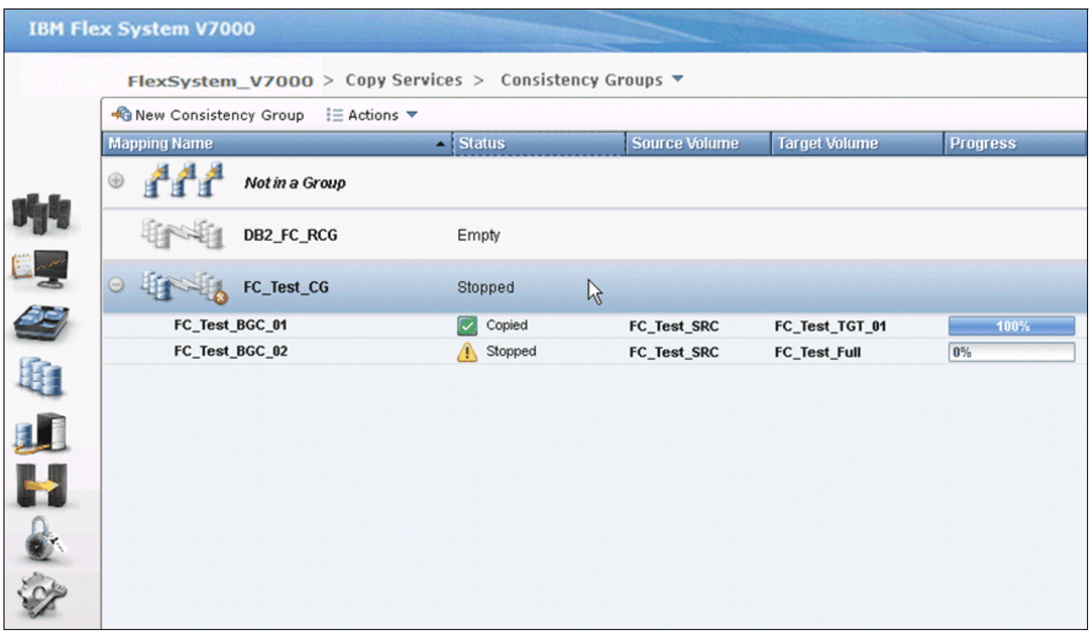


Figure 9-52 Consistency group stop completes

Previously copied relationships that were added to a consistency group that was later stopped before all members of the consistency group completed synchronization do not go out of the Copied state.

Removing FlashCopy mappings from a consistency group

The FlashCopy mappings can be removed from a consistency group by selecting the FlashCopy mappings and clicking **Remove from Consistency Group** from the **Actions** drop-down menu of the FlashCopy mapping (Figure 9-53).

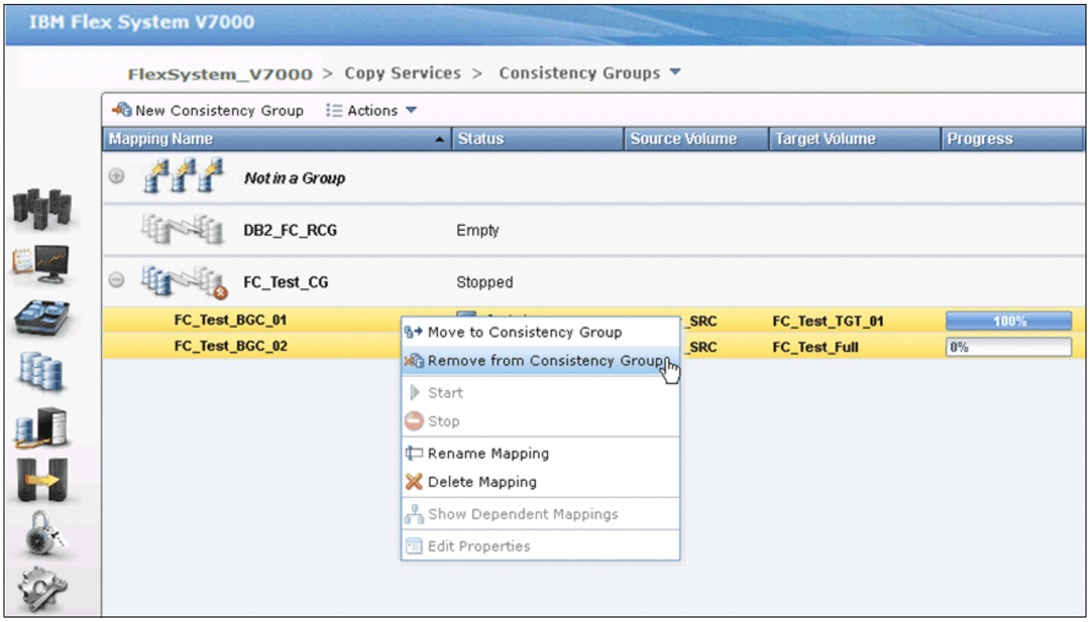


Figure 9-53 Remove from consistency group

The FlashCopy mappings are returned to the Not in a Group list after being removed from the consistency group.

Deleting a consistency group

A consistency group can be deleted by clicking **Delete** from the **Actions** drop-down menu (Figure 9-54).

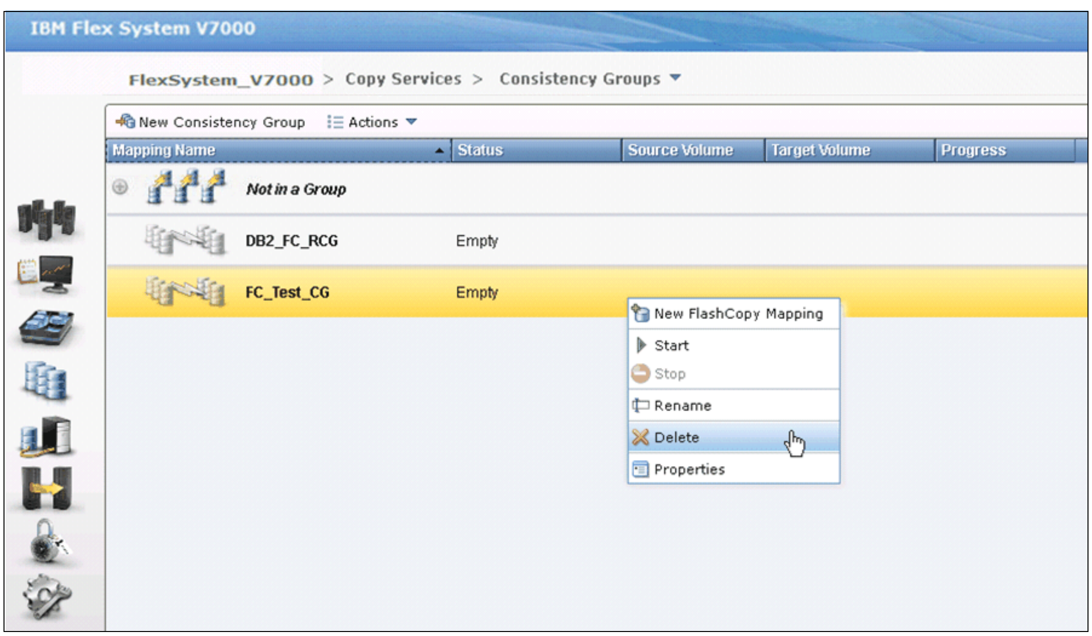


Figure 9-54 Delete a consistency group

9.3 Remote Copy

In this section, we describe how the Remote Copy function works in IBM Flex System V7000 Storage Node. We also provide the implementation steps, using the GUI, for Remote Copy configuration and management.

Remote Copy consists of three methods for copying: Metro Mirror, Global Mirror, and Global Mirror with Change Volumes. Metro Mirror is designed for metropolitan distances with a synchronous copy requirement. Global Mirror is designed for longer distances without requiring the hosts to wait for the full round-trip delay of the long-distance link through asynchronous methodology. With Version 6.4.1, Global Mirror with Change Volumes is an added piece of functionality for Global Mirror designed to attain consistency on lower-quality network links.

Metro Mirror and Global Mirror are IBM branded terms for the functions Synchronous Remote Copy and Asynchronous Remote Copy. Throughout this book, the term “Remote Copy” is used to refer to both functions where the text applies to each term equally.

9.3.1 Remote Copy concepts

In the remote copy there are a number of different possible concepts that may be used when performing the desired functions to successfully reach the desired goal. In this session we discuss these different concepts and the reasons for and results of their use.

Partnership

When creating a partnership, connect an IBM Flex System V7000 Storage Node to either another IBM Flex System V7000 Storage Node, IBM Storwize V7000 systems or IBM SAN Volume Controller (SVC) systems which are in separate clusters. These clusters may be separated by distance, but distance is not a requirement. After the partnership creation has been configured on both systems, further communication between the node canisters in each of the storage systems is established and maintained by the SAN network. All inter-cluster communication goes over either a Fibre Channel network or over an FCoE network. Partnership must be defined on both storage systems to make the partnership fully functional.

Though partnerships for remote copy are created between the IBM virtual storage systems, external storage system which are virtualized behind an IBM virtual storage system can be used as members of the storage pools which have source or target volumes as members of a remote mirrored pair. This does not mean that the external storage’s mirroring product is supported; but that the mdisk members can be accessed as storage members of the IBM virtual storage system for use as virtual storage locations. This factor can impact the licensing of the remote copy service on the IBM virtual storage systems.

A parameter named layer has been added to the clustered system properties which can be changed from storage (default) to replication if you need to make a new relationship between the IBM Flex System V7000 Storage Node and an SVC system. You can only change this parameter using the CLI command chsystem. In SVC systems, this parameter is fixed to replication and cannot be changed.

> **Interconnection:** Interconnects between an IBM Flex System V7000 Storage Node, an IBM Storwize V7000 or IBM SAN Volume Controller requires the use of Version 6.4.1. Both systems in the partnership must be at that level, and the IBM Flex System V7000 Storage Node must be set to the *replication layer* using the `svctask chsystem -layer replication`. Usage limitations are described in “Introduction to layers”.

Introduction to layers

Version 6.4.1 uses a concept called *layers* for the IBM Flex System V7000 Storage Node. The setting for the *layers* determine how the IBM Flex System V7000 Storage Node interacts with the other IBM Flex System V7000 Storage Node, IBM Storwize V7000 or IBM SAN Volume Controller. Currently there are two layers: replication and storage.

The *replication layer* is used when you want to use the IBM Flex System V7000 Storage Node with one or more other IBM Flex System V7000 Storage Node, IBM Storwize V7000 and a SAN Volume Controller is also a Remote Copy partner. The *storage layer* is the default mode of operation for the IBM Flex System V7000 Storage Node and IBM Storwize V7000, and is used when you want to only remote copy from the IBM Flex System V7000 Storage Node to another IBM Flex System V7000 Storage Node or an IBM Storwize V7000. It is also used when presenting storage to the IBM Flex System V7000 Storage Node from an IBM Storwize V7000 as external storage for it to manage.

> **Note:** Other IBM or 3rd party storage devices licensed and virtualized behind the IBM SVC may be included in the copy services.

The layer for the IBM Flex System V7000 Storage Node can be switched by running **`svctask chsystem -layer replication`**. Generally, switch the layer while your IBM Flex System V7000 Storage Node is not in production. This situation prevents potential disruptions, as layer changes are not I/O tolerant.

Figure 9-55 shows the effect of layers on IBM Flex System V7000 Storage Node, IBM Storwize V7000 and SAN Volume Controller partnerships.

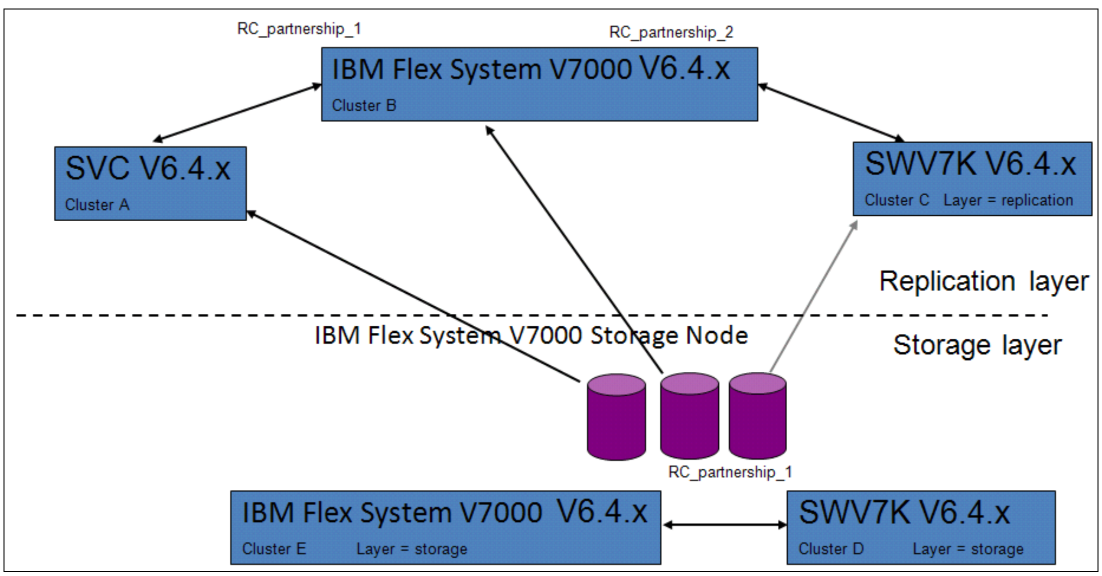


Figure 9-55 The relationship of layers and SAN Volume Controller and Storwize V7000 partnerships

The *replication layer* allows an IBM Flex System V7000 Storage Node to be Remote Copy partners with a SAN Volume Controller. The *storage layer* allows an IBM Storwize V7000 to Remote Copy to an IBM Flex System V7000 Storage Node which is also configured to *storage layer,* or to function as external back-end storage for an IBM Flex System V7000 Storage Node which is configured to *replication layer*. An IBM Flex System V7000 Storage Node cannot be in both layers at the same time.

Limitations on the SAN Volume Controller and Storwize V7000 partnership

With Version 6.4.1, the IBM Flex System V7000 Storage Node, SAN Volume Controller, and IBM Storwize V7000 systems can be partners in a Remote Copy partnership. There are however, a few limitations:

- The layer for both the IBM Flex System V7000 Storage Node and the IBM Storwize V7000 must be set to replication for remote copy to work with the SVC in the mix. The default is *storage*.
- If any other IBM Flex System V7000 Storage Node, SAN Volume Controller or IBM Storwize V7000 ports are visible on the SAN, aside from the ones on the cluster where you are making changes, you cannot change the layer.
- If any host object is defined to an IBM Flex System V7000 Storage Node, SAN Volume Controller or IBM Storwize V7000 system attached, you cannot change the layer.
- If any MDisks from an IBM Flex System V7000 Storage Node or IBM Storwize V7000 other than the one you are making the layer change on are visible, you cannot change the layer.
- If any cluster partnership is defined, you cannot change the layer.

Partnership topologies

A partnership between up to four IBM Flex System V7000 Storage Node are allowed.

Typical partnership topologies between multiple IBM Flex System V7000 Storage Nodes are:

- Daisy-chain topology (Figure 9-56)

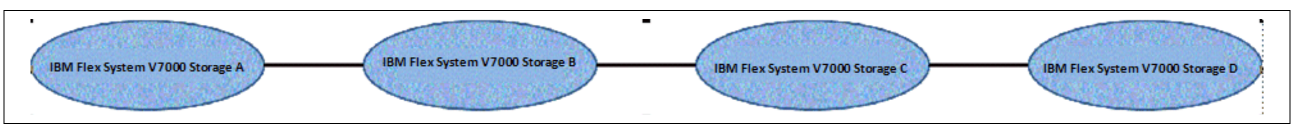


Figure 9-56 Daisy-chain topology for IBM Flex System V7000 Storage Node

- Triangle topology (Figure 9-57)

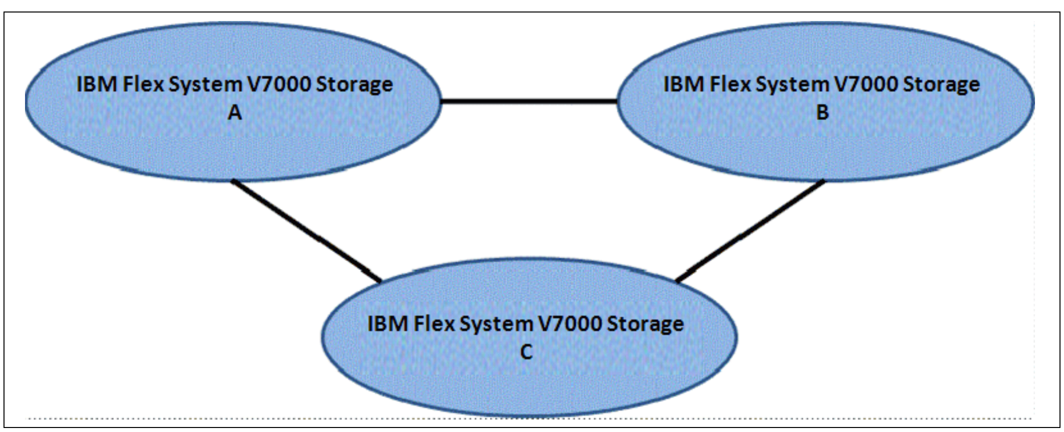


Figure 9-57 Triangle topology for IBM Flex System V7000 Storage Node

- Star topology (Figure 9-58 on page 383)

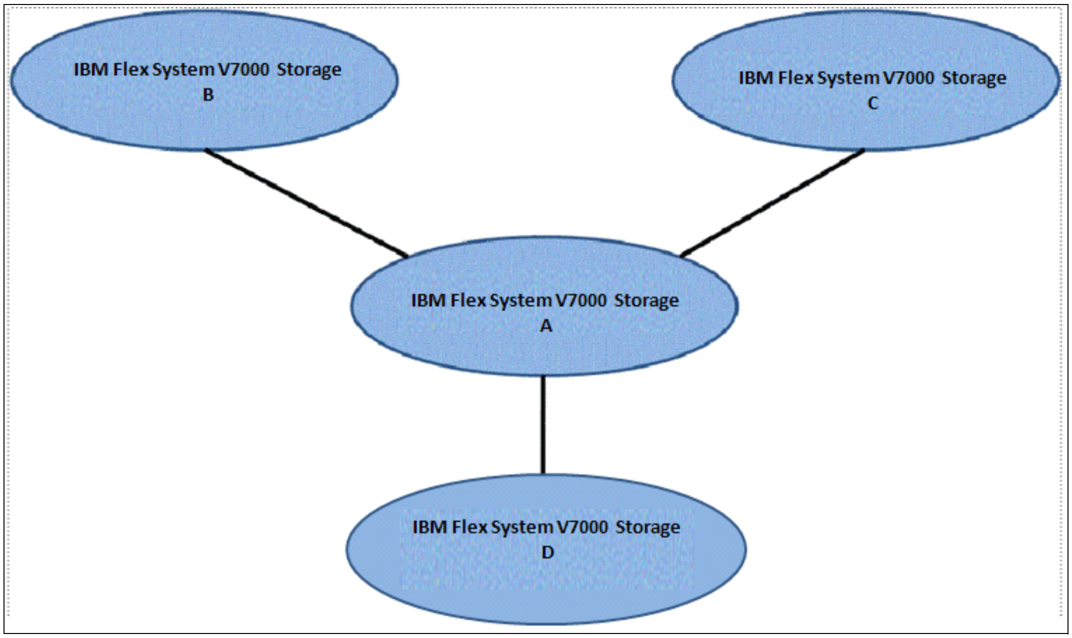


Figure 9-58 Star topology for IBM Flex System V7000 Storage Node

- Full-meshed topology (Figure 9-59)

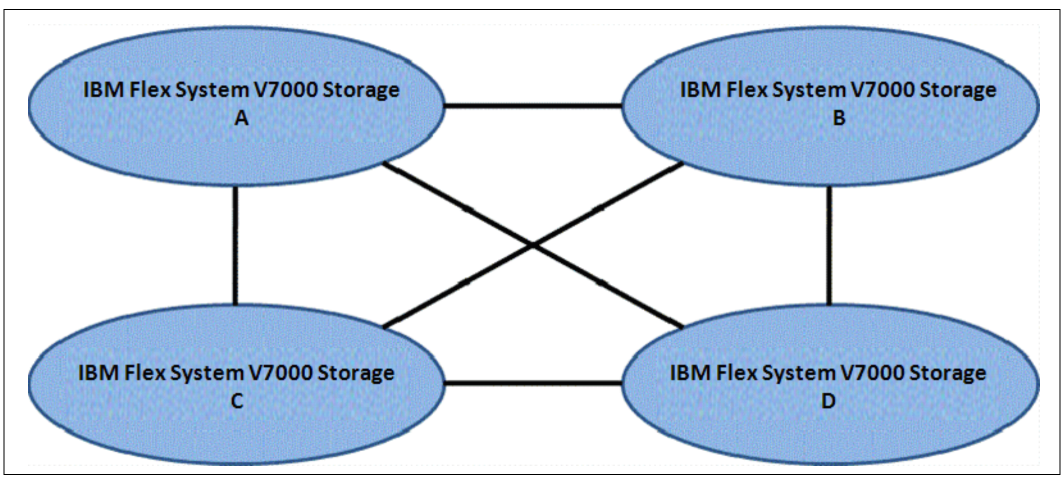


Figure 9-59 Full-meshed IBM Flex System V7000 Storage Node

Partnerships: These partnerships are also valid for mixed configurations with SAN Volume Controllers as long as the replication layer is set.

Partnership states

A partnership has the following states:

- Partially Configured

  Indicates that only one cluster partner is defined from a local or remote cluster to the displayed cluster and is started. For the displayed cluster to be configured fully and to complete the partnership, you must define the cluster partnership from the cluster that is displayed to the corresponding local or remote cluster.

- Fully Configured

Indicates that the partnership is defined on the local and remote clusters and is started.

- Remote Not Present

  Indicates that the remote cluster is not present for the partnership.

- Partially Configured (Local Stopped)

  Indicates that the local cluster is only defined to the remote cluster and the local cluster is stopped.

- Fully Configured (Local Stopped)

  Indicates that a partnership is defined on both the local and remote clusters and the remote cluster is present, but the local cluster is stopped.

- Fully Configured (Remote Stopped)

  Indicates that a partnership is defined on both the local and remote clusters and the remote cluster is present, but the remote cluster is stopped.

- Fully Configured (Local Excluded)

  Indicates that a partnership is defined between a local and remote cluster; however, the local cluster has been excluded. Usually this state occurs when the fabric link between the two clusters has been compromised by too many fabric errors or slow response times of the cluster partnership.

- Fully Configured (Remote Excluded)

  Indicates that a partnership is defined between a local and remote cluster; however, the remote cluster has been excluded. Usually this state occurs when the fabric link between the two clusters has been compromised by too many fabric errors or slow response times of the cluster partnership.

- Fully Configured (Remote Exceeded)

  Indicates that a partnership is defined between a local and remote cluster and the remote is available; however, the remote cluster exceeds the number of allowed clusters within a cluster network. The maximum of four clusters can be defined in a network. If the number of clusters exceeds that limit, the IBM Flex System V7000 Storage Node determines the inactive cluster or clusters by sorting all the clusters by their unique identifier in numerical order. The inactive cluster partner that is not in the top four of the cluster unique identifiers shows Fully Configured (Remote Exceeded).

Remote Copy relationships

A Remote Copy relationship is a relationship between two individual volumes of the same size. These volumes are called a *master (source) volume* and an *auxiliary (target) volume*.

Typically, the master volume contains the production copy of the data and is the volume that the application normally accesses. The auxiliary volume typically contains a backup copy of the data and is used for disaster recovery.

The master and auxiliary volumes are defined when the relationship is created, and these attributes never change. However, either volume can operate in the primary or secondary role as necessary. The primary volume contains a valid copy of the application data and receives updates from the host application, which is analogous to a source volume. The secondary volume receives a copy of any updates to the primary volume, because these updates are all transmitted across the mirror link. Therefore, the secondary volume is analogous to a continuously updated target volume. When a relationship is created, the master volume is assigned the role of primary volume and the auxiliary volume is assigned the role of secondary volume. The initial copying direction is from master to auxiliary. When the relationship is in a consistent state, you can reverse the copy direction.

The two volumes in a relationship must be the same size. The Remote Copy relationship can be established on the volumes within one IBM Flex System V7000 Storage Node system, which is called an *intra*-cluster relationship. The relationship can also be established between two IBM Flex System V7000 Storage Nodes, or with an IBM Storwize V7000 storage systems or an IBM SAN Volume Controller, which are called *inter*-cluster relationships.

Usage of Remote Copy target volumes as Remote Copy source volumes is not allowed. A FlashCopy target volume can be used as Remote Copy source volume and also as a Remote Copy target volume.

FlashCopy: Using target volumes as remote copy sources and target volumes for FlashCopy requires Version 6.2.x or higher.

Metro Mirror

Metro Mirror is a type of Remote Copy that creates a synchronous copy of data from a master volume to an auxiliary volume. With synchronous copies, host applications write to the master volume, but do not receive confirmation that the write operation has completed until the data is written to the auxiliary volume. This action ensures that both volumes have identical data when the copy completes. After the initial copy completes, the Metro Mirror function maintains a fully synchronized copy of the source data at the target site at all times.

Figure 9-60 shows how a write to the master volume is mirrored to the cache of the auxiliary volume before an acknowledgement of the write is sent back to the host that issued the write. This process ensures that the auxiliary is synchronized in real time, if it is needed in a failover situation.

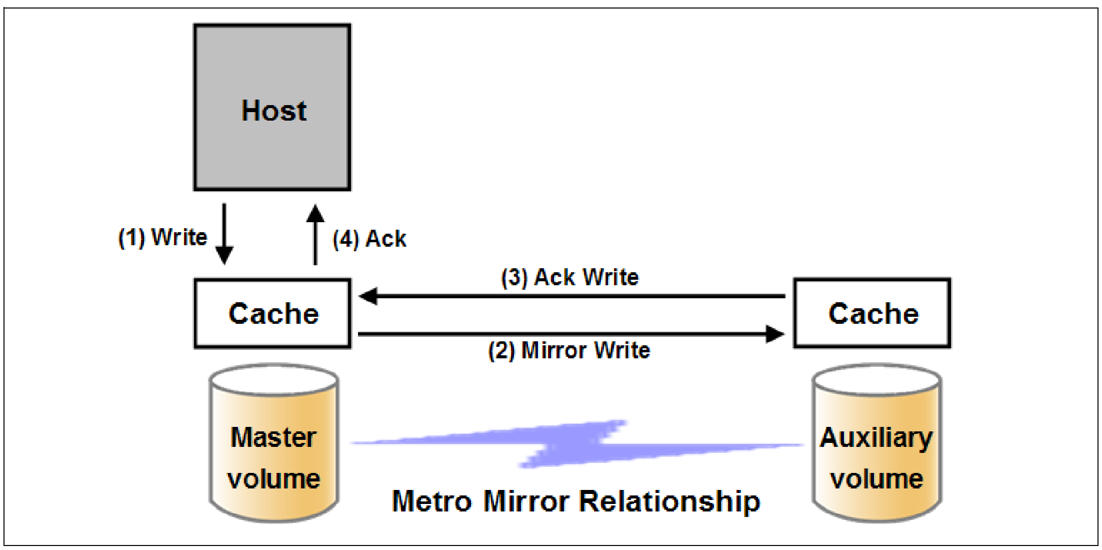


Figure 9-60 Write on volume in a Metro Mirror relationship

The Metro Mirror function supports copy operations between volumes that are separated by distances up to 300 km. For disaster recovery purposes, Metro Mirror provides the simplest way to maintain an identical copy on both the primary and secondary volumes. However, as with all synchronous copies over remote distances, there can be a performance impact to host applications. This performance impact is related to the distance between primary and secondary volumes and, depending on application requirements, its use might be limited based on the distance between sites.

Global Mirror

Global Mirror provides an asynchronous copy, which means that the secondary volume is not an exact match of the primary volume at every point in time. The Global Mirror function provides the same function as Metro Mirror Remote Copy without requiring the hosts to wait for the full round-trip delay of the long-distance link; however, some delay can be seen on the hosts in congested or overloaded environments. Make sure that you closely monitor and understand your workload.

In asynchronous Remote Copy, which Global Mirror provides, write operations are completed on the primary site and the write acknowledgement is sent to the host before it is received at the secondary site. An update of this write operation is sent to the secondary site at a later stage, which provides the capability to perform Remote Copy over distances exceeding the limitations of synchronous Remote Copy.

The distance of Global Mirror replication is limited primarily by the latency of the WAN Link provided. Global Mirror has a requirement of 80 ms round-trip-time for data sent to the remote location. The propagation delay is roughly 8.2 µs per mile or 5 µs per kilometer for Fibre Channel connections. Each device in the path adds addition delay of about 25 µs. Devices that use software (such as some compression devices) add much more time. The time added by software-assisted devices is highly variable and should be measured directly. Be sure to include these times when planning your Global Mirror design.

You should also measure application performance based on the expected delays before Global Mirror is fully implemented. The IBM Flex System V7000 Storage Node provides you with an advanced feature of Global Mirror that permits you to test performance implications before deploying Global Mirror and obtaining a long-distance link. This advanced feature is enabled by modifying the IBM Flex System V7000 Storage Node parameters **`gmintradelaysimulation`** and **`gminterdelaysimulation`**. These two parameters could be used to simulate the write delay to the secondary volume. The delay simulation can be enabled separately for each intra-cluster or inter-cluster Global Mirror. You can use this feature to test an application before the full deployment of the Global Mirror feature.

Figure 9-61 on page 387 shows that a write operation to the master volume is acknowledged back to the host issuing the write before the write operation is mirrored to the cache for the auxiliary volume.

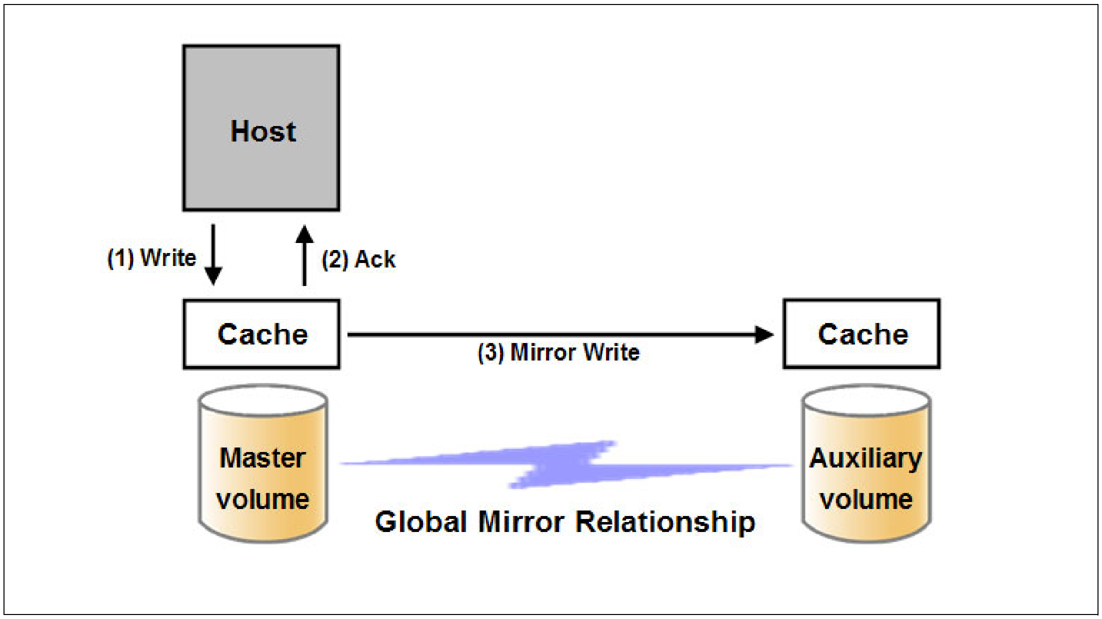


Figure 9-61 Global Mirror write sequence

The Global Mirror algorithms maintain a consistent image on the auxiliary volume at all times. They achieve this consistent image by identifying sets of I/Os that are active concurrently at the master, assigning an order to those sets, and applying those sets of I/Os in the assigned order at the secondary.

In a failover scenario, where the secondary site needs to become the master source of data, depending on the workload pattern and the bandwidth and distance between local and remote site, certain updates might be missing at the secondary site. Therefore, any applications that use this data must have an external mechanism for recovering the missing updates and reapplying them, for example, a transaction log replay.

9.3.2 Global Mirror with Change Volumes

Global Mirror within the IBM Flex System V7000 Storage Node is designed to achieve a recovery point objective (RPO) which is as low as possible, so that data is as up-to-date as possible. This capability places some strict requirements on your infrastructure and in certain situations, with low network link quality or congested or overloaded hosts, you maybe impacted by multiple 1920 (congestion) errors.

Congestion errors happen in three primary situations:

1. Congestion at the source site through the host or network
2. Congestion on the network link or network path
3. Congestion at the target site through the host or network

With all releases after 6.3.0, Global Mirror picked up new functionality that is designed to address a few conditions that were negatively effecting some Global Mirror implementations:

- Estimation of bandwidth requirements tends to be complex.
- It is often difficult to guarantee the latency and bandwidth requirements can be met.
- Congested hosts on either the source or target site can cause disruption.
- Congested network links can cause disruption with only intermittent peaks.

To address these issues, *Change Volumes* were added as an option for Global Mirror relationships. Change Volumes use the FlashCopy functionality, but cannot be manipulated as FlashCopy volumes, as they are special purpose only. Change volumes are used at both the primary and secondary end of the replication link. At the primary the change volume stores copy-on-write changes from primary that may need to be sent to secondary to preserve the point-in-time when cycling period last started. At the secondary the change volume is used to ensure a consistent point-in-time image for recovery if a disaster recovery needs to be implemented. Change Volumes replicate point-in-time images on a cycling period (the default is 300 seconds.) This situation means that your change rate only needs to include the condition of the data at the point-in-time the image was taken, instead of all the updates during the period. This situation can provide significant reductions in replication volume.

Figure 9-62 shows a diagram of a basic Global Mirror relationship without Change Volumes.

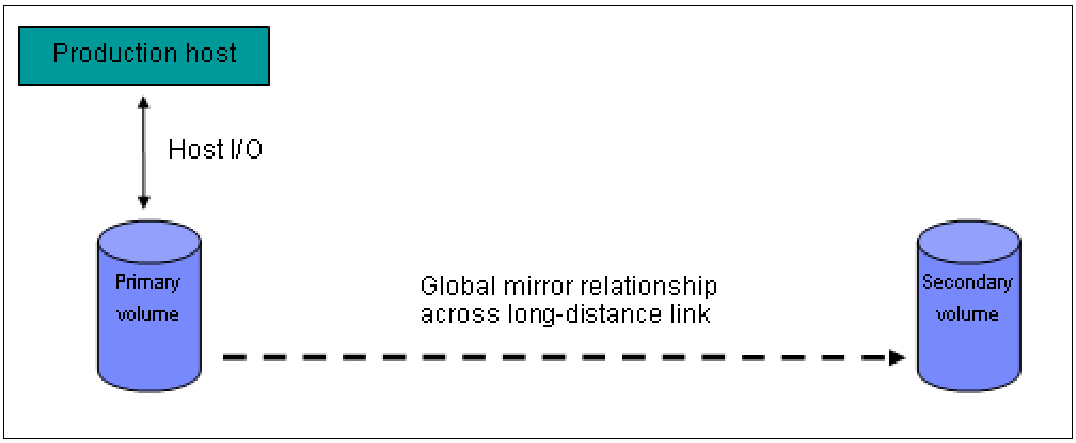


Figure 9-62 Global Mirror without Change Volumes

Figure 9-63 shows a diagram with the Change Volumes.

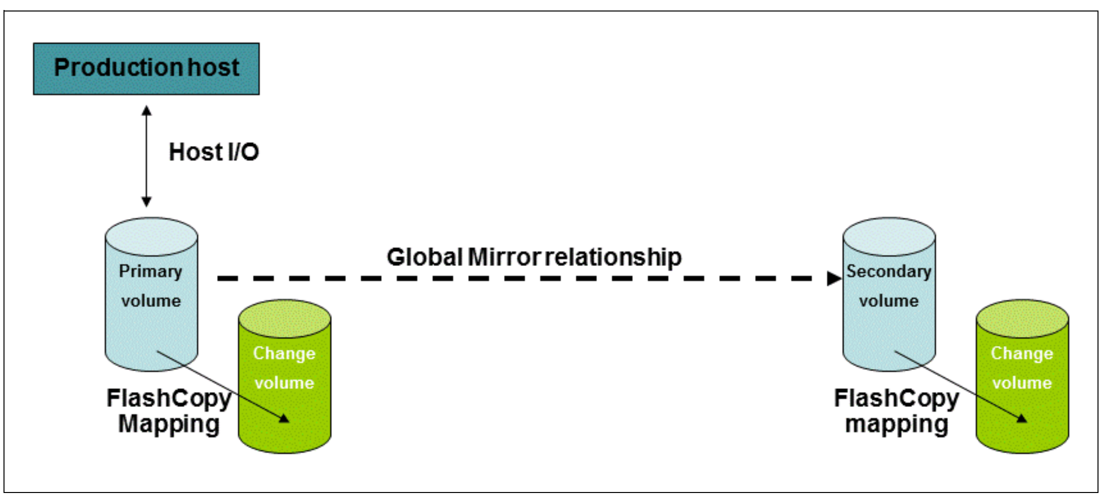


Figure 9-63 Global Mirror with Change Volumes

With Change Volumes, a FlashCopy mapping exists between the primary volume and the primary Change Volume. The mapping is updated during a cycling period (every 60 seconds to one day.) The primary Change Volume is then replicated to the secondary Global Mirror volume at the target site, which is then captured in another change volume on the target site.

This situation provides an always consistent image at the target site and protects your data from being inconsistent during re synchronization.

Take a closer look at how Change Volumes might reduce replication traffic.

Figure 9-64 shows a number of I/Os on the source volume and the same number on the target volume, and in the same order. Assuming that this set is the same set of data being updated over and over, then these updates are wasted network traffic and the I/O can be completed much more efficiently (Figure 9-65).

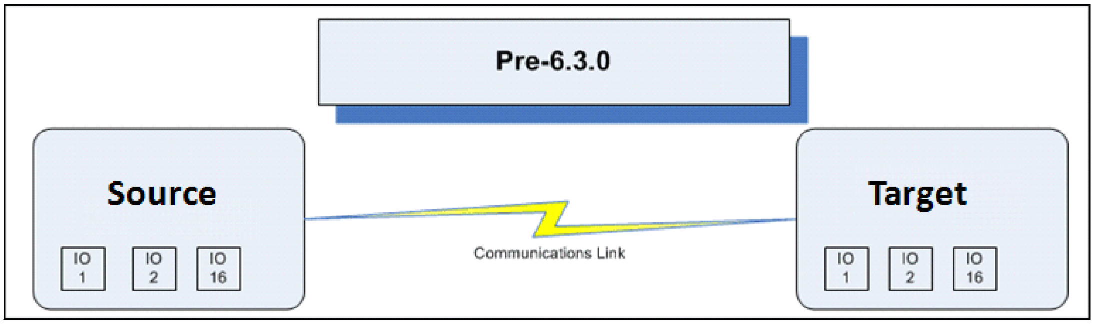


Figure 9-64 Global Mirror I/O replication without Change Volumes

In Figure 9-65, the same data is being updated repeatedly, so Change Volumes demonstrate significant IO transmission savings, because you only need to send I/O number 16, which was the last I/O before the cycling period.

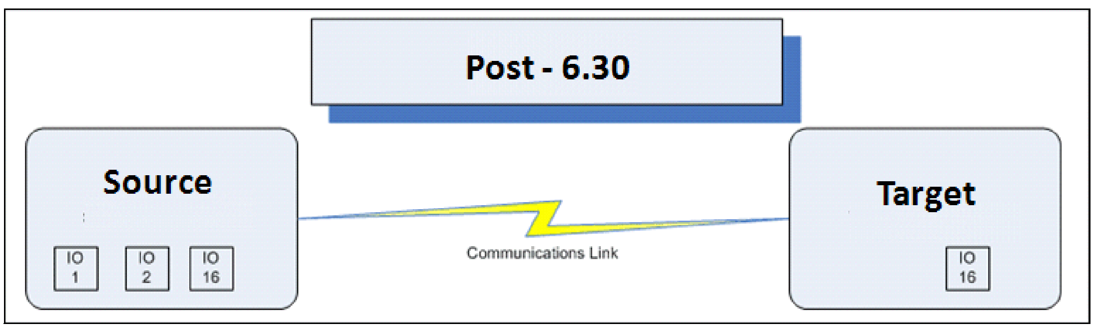


Figure 9-65 Global Mirror I/O replication with Change Volumes

The cycling period can be adjusted by running **`chrcrelationship -cycleperiodseconds <60-86400>`**. If a copy does not complete in the cycle period, the next cycle does not start until the prior one has completed. It is for this reason that using Change Volumes gives you two possibilities for RPO:

- If your replication completes in the cycling period, then your RPO is twice the cycling period.
- If your replication does not complete within the cycling period, then your RPO is twice the completion time. The next cycling period starts immediately after the prior one is finished.

Careful consideration should be put into balancing your business requirements with the performance of Global Mirror with Change Volumes. Global Mirror with Change Volumes increases the inter-cluster traffic for more frequent cycling periods, so going as short as possible is not always the answer. In most cases, the default should meet your requirements and perform reasonably well.

Important: When using Global Mirror volumes with Change Volumes, make sure that you remember to select the Change Volume on the auxiliary (target) site. Failure to do so leaves you exposed during a resynchronization operation.

Important: The GUI for Version 6.4.1 automatically creates Change Volumes for you. However, it is a limitation of this initial release that they are fully provisioned volumes. To save space, you should create thin-provisioned volumes before and use the existing volume option to select your change volumes.

Remote Copy consistency groups

A consistency group is a logical entity that groups copy relationships. By grouping the relationships, you can ensure that these relationships are managed in unison and the data within the group is in a consistent state. More information about the necessity of consistency groups can be found in "FlashCopy consistency groups" on page 348.

Remote Copy commands can be issued to a Remote Copy consistency group, and therefore simultaneously for all Metro Mirror relationships defined within that consistency group, or to a single Metro Mirror relationship that is not part of a Metro Mirror consistency group.

Figure 9-66 shows the concept of Remote Copy consistency groups. Because the RC_Relationships 1 and 2 are part of the consistency group, they can be handled as one entity, while the stand-alone RC_Relationship 3 is handled separately.

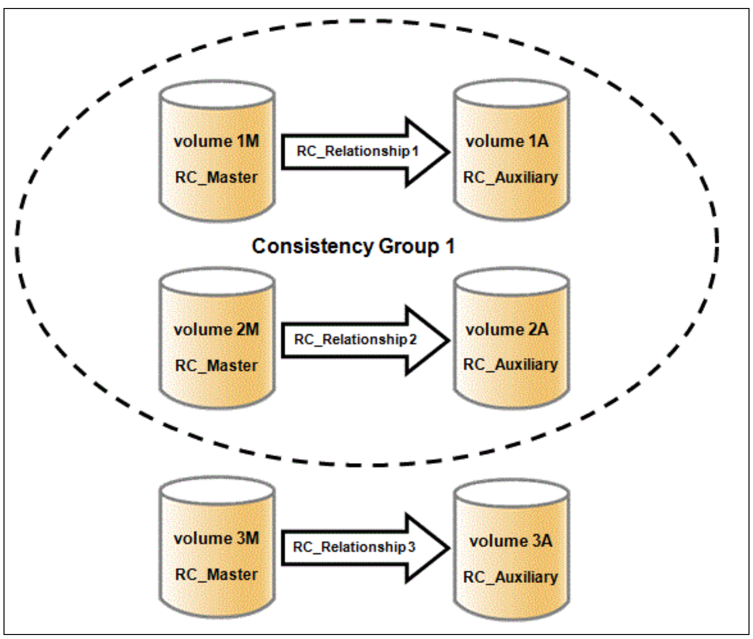


Figure 9-66 Remote Copy consistency group

Remote Copy relationships can only belong to one consistency group, but they do not have to belong to a consistency group. Relationships that are not part of a consistency group are called *stand-alone relationships*. A consistency group can contain zero or more relationships. All relationships in a consistency group must have matching primary and secondary clusters, which are sometimes referred to as *master* and *auxiliary clusters*. All relationships in a consistency group must also have the same copy direction and state.

Metro Mirror and Global Mirror relationships cannot belong to the same consistency group. A copy type is automatically assigned to a consistency group when the first relationship is

added to the consistency group. After the consistency group is assigned a copy type, only relationships of that copy type can be added to the consistency group.

Remote Copy and consistency group states

Stand-alone Remote Copy relationships and consistency groups share a common configuration and state model. All of the relationships in a non-empty consistency group have the same state as the consistency group.

These states apply to both the relationships and the consistency groups, except for the Empty state, which is only for consistency groups:

- InconsistentStopped

  The primary volumes are accessible for read and write I/O operations, but the secondary volumes are not accessible for either one. A copy process must be started to make the secondary volumes consistent.

- InconsistentCopying

  The primary volumes are accessible for read and write I/O operations, but the secondary volumes are not accessible for either one. This state indicates that a copy process is ongoing from the primary to the secondary volume.

- ConsistentStopped

  The secondary volumes contain a consistent image, but it might be out-of-date with respect to the primary volumes. This state can occur when a relationship was in the ConsistentSynchronized state and experiences an error that forces a freeze of the consistency group or the Remote Copy relationship.

- ConsistentSynchronized

  The primary volumes are accessible for read and write I/O operations. The secondary volumes are accessible for read-only I/O operations.

- Idling

  Both the primary volumes and the secondary volumes are operating in the primary role. Consequently, the volumes are accessible for write I/O operations.

- IdlingDisconnected

  The volumes in this half of the consistency group are all operating in the primary role and can accept read or write I/O operations.

- InconsistentDisconnected

  The volumes in this half of the consistency group are all operating in the secondary role and cannot accept read or write I/O operations.

- ConsistentDisconnected

  The volumes in this half of the consistency group are all operating in the secondary role and can accept read I/O operations but not write I/O operations.

- Empty

  The consistency group does not contain any relationships.

9.3.3 Remote Copy planning

Before you use Remote Copy, you need to plan for its usage.

General guidelines for Remote Copy

General guidelines for Remote Copy include the following items:

- Partnerships between a mix of up to four IBM Flex System V7000 Storage Node, IBM Storwize V7000 storage systems or IBM SAN Volume Controller systems is allowed. The partnership must be defined on all partnered members to make it fully functional.
- The two volumes in a relationship must be the same size.
- The Remote Copy relationship can be established on the volumes within one IBM Flex System V7000 Storage Node system or in different IBM Flex System V7000 Storage Node system. When the two volumes are in the same cluster, they must be in the same I/O group.
- You can use Remote Copy target volumes as Remote Copy source volumes. A FlashCopy target volume can be used as Remote Copy source volume, as of Version 6.2.0. There are additional restrictions outlined in Table 9-5 on page 396.
- The Metro Mirror function supports copy operations between volumes that are separatedby distances up to 300 km. Metro Mirroring is supported over either an FC or FCoE SAN infrastructure.
- Global Mirroring function supports copy operations between volumes that are separated by distances over 300 km. Global Mirroring is only supported over an FC SAN infrastructure.

  At this time no remote mirroring functions are supported over the iSCSI fabric. Consult IBM System Storage Interoperation Center (SSIC) for supported configurations at:
  `http://www-03.ibm.com/systems/support/storage/ssic/interoperability.wss`
- One Remote Copy relationship can only belong to one consistency group.
- All relationships in a consistency group must have matching primary and secondary clusters, which are sometimes referred to as master and auxiliary clusters. All relationships in a consistency group must also have the same copy direction and state.
- Metro Mirror and Global Mirror relationships cannot belong to the same consistency group.
- To manage multiple Remote Copy relationships as one entity, relationships can be made part of a Remote Copy consistency group, which ensures data consistency across multiple Remote Copy relationships and provides ease of management.
- An IBM Flex System V7000 storage system implements flexible resynchronization support, enabling it to resynchronize volume pairs that have experienced write I/Os to both disks and to resynchronize only those regions that are known to have changed.
- With IBM Flex System V7000, customers may wish to use the global mirror capability which enables them to have a low bandwidth link between sites. This capability is known

as Global Mirror with Change Volumes and should have Change Volumes defined for both the source and target volumes when used.

This capability provides options to help administrators balance network bandwidth requirements and RPO times for applications, helping reduce operation costs for disaster recovery solutions. Remote mirroring supports higher RPO times by allowing the data at the disaster recovery site to get further out of sync with the production site if the communication link limits replication, and then approaches synchronicity again when the link is not as busy. This low bandwidth remote mirroring uses space-efficient FlashCopy targets as sources in remote copy relationships to increase the time allowed to complete a remote copy data cycle.

Licensing guidelines for Remote Copy

To authorize use of remote copy capabilities of the IBM Flex System V7000, you must purchase a feature license for IBM Flex System V7000 Remote Mirroring Software. This license is based on the quantity of enclosures that are authorized for the feature.

To determine the number of authorized enclosures needed you should count all enclosures on the system that are managed by the IBM Flex System V7000 Disk System, including each internal enclosure licensed with the IBM Flex System V7000 Base Software and each external enclosure licensed with the IBM Flex System V7000 External Virtualization Software. This should be done at both the ends of the mirroring relationship. The location with the least number of members is the location that defines the number of licenses needed for each location site.

For example, if you have site A with a DS5020 consisting of three enclosures managed by an IBM Flex System V7000 consisting of one control enclosure and one expansion enclosure mirroring to a partner at site B with an IBM Flex System V7000 consisting of one control enclosure and two expansion enclosures and no external virtual storage system attached, then in order to authorize remote mirroring for this configuration you will need a license to the IBM Flex System V7000 Remote Mirroring Software, for a quantity of six enclosures; three for site A and three for site B. If however, in the future you were to grow your site B system to include more enclosures or an external virtualized system you would then need to re-evaluate the licensing to add the appropriate number for the new configuration of the site with the fewest enclosures managed.

Remote mirroring Flex System V7000 with an SVC in the mix, because of a difference in metrics, SVC mirroring can be licensed for a subset of the total storage virtualized. However, the IBM Flex System V7000 and the IBM Storwize V7000 still must used the above model for their licensing.

Remote Copy configuration limits

Table 9-4 lists the Remote Copy configuration limits.

Table 9-4 Remote Copy configuration limits

Parameter	Value
Number of Remote Copy consistency groups per cluster	256
Number of Remote Copy relationships per consistency group	8,192
Number of Remote Copy relationships per I/O Group	2,048

Parameter	Value
Total Remote Copy volume capacity per I/O Group	1024 TB (This limit is the total capacity for all master and auxiliary volumes in the I/O group.)

SAN planning for Remote Copy

Here are guidelines for planning for a SAN for Remote Copy.

Zoning recommendation

Node canister ports on each IBM Storwize V7000 must be able to communicate with each other in order for the partnership creation to be performed. These ports must be visible to each other on your SAN. Proper switch zoning is critical to facilitating inter-cluster communication.

The SAN zoning recommendation is as follows:

- For each node canister, exactly two Fibre Channel ports should be zoned to exactly two Fibre Channel ports from each node canister in the partner cluster.
- If dual-redundant inter-switch links (ISLs) are available, then the two ports from each node should be split evenly between the two ISLs, that is, exactly one port from each node canister should be zoned across each ISL. More details are available at the following address:

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003634&myns=s033&mynp=familyind5329743&mync=E
- Additionally, all local zoning rules should be followed. A properly configured SAN fabric is key to not only local SAN performance, but Remote Copy. You can learn more about these rules at the following address:

  http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp?topic=%2Fcom.ibm.storwize.v7000.doc%2Fsvc_configrulessummary_02171530.html

> **Fabrics:** When a local fabric and a remote fabric are connected together for Remote Copy purposes, the ISL hop count between a local node and a remote node cannot exceed seven.

Remote Copy link requirements

The link requirements are valid for Metro Mirror and Global Mirror. Basically, there are two link requirements:

- Round-trip latency

  The total round-trip latency must be less than 80 ms, and less than 40 ms in each direction. Latency simulations should be performed with your applications before putting any network links in place to see if the applications perform at an acceptable level while meeting the round-trip latency requirement.

- Bandwidth

  The bandwidth must satisfy the following requirements:

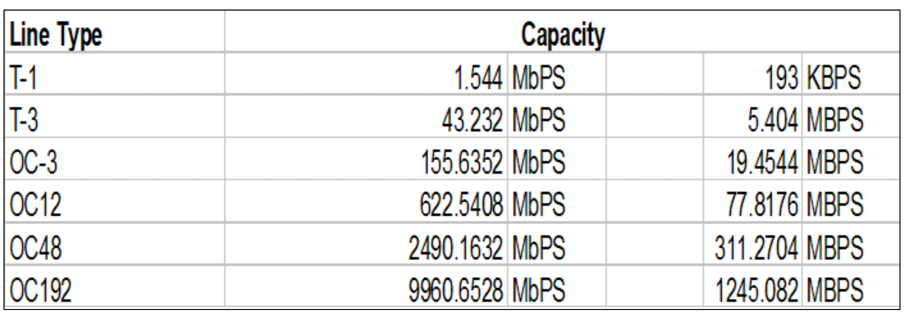
  – If you are not using Change Volumes: Be able to sustain peak write load for all mirrored volumes and background copy traffic.

  – If you are using Change Volumes with Global Mirror: Be able to sustain change rate of Source Change Volumes and background copy traffic.

  – Additional background copy rate (the best practice is 10% to 20% of maximum peak load) for initial synchronization and resynchronization.

  – Remote Copy internal communication at idle with or without Change Volumes is approximately 2.6 Mbps. This amount is the minimum amount.

  Figure 9-67 shows a table of the most common bandwidths available for wide area network (WAN) line types available in the industry.



Line Type	Capacity				
T-1	1.544	MbPS		193	KBPS
T-3	43.232	MbPS		5.404	MBPS
OC-3	155.6352	MbPS		19.4544	MBPS
OC12	622.5408	MbPS		77.8176	MBPS
OC48	2490.1632	MbPS		311.2704	MBPS
OC192	9960.6528	MbPS		1245.082	MBPS

Figure 9-67 Wide area network bandwidth speed and feed chart

Redundancy: If the link between two sites is configured with redundancy so that it can tolerate single failures, the link must be sized so that the bandwidth and latency requirement can be met during single failure conditions.

Interaction between Remote Copy and FlashCopy

Table 9-5 on page 396 lists which combinations of FlashCopy and Remote Copy are supported.

Table 9-5 FlashCopy and Remote Copy interaction

Component	Remote Copy primary site	Remote Copy secondary site
FlashCopy source	Supported	Supported. When the FlashCopy relationship is in the Preparing and Prepared states, the cache at the Remote Copy secondary site operates in write-through mode. This process adds additional latency to the already latent Remote Copy relationship.
FlashCopy destination	This combination is supported. It has several restrictions: ► Issuing `stop -force` might cause the Remote Copy relationship to fully resynchronize. ► The code level must be Version 6.2.x or higher. ► The I/O group must be the same.	This combination is supported with the restriction that the FlashCopy mapping cannot be copying, stopping, or suspended. Otherwise, the restrictions are the same as at the Remote Copy primary site.

If you are not using Global Mirror with Change Volumes, you can, for disaster recovery purposes, use the FlashCopy feature to create a consistent copy of an image before you restart a Global Mirror relationship.

When a consistent relationship is stopped, the relationship enters the consistent_stopped state. While in this state, I/O operations at the primary site continue to run. However, updates are not copied to the secondary site. When the relationship is restarted, the synchronization process for new data is started. During this process, the relationship is in the inconsistent_copying state. The secondary volume for the relationship cannot be used until the copy process completes and the relationship returns to the consistent state. When this situation occurs, start a FlashCopy operation for the secondary volume before you restart the relationship. While the relationship is in the Copying state, the FlashCopy feature can provide a consistent copy of the data. If the relationship does not reach the synchronized state, you can use the FlashCopy target volume at the secondary site.

9.4 Troubleshooting Remote Copy

Remote Copy (Global Mirror and Metro Mirror) has two primary error codes: 1920 or 1720. A 1920 error is a congestion error. This error means that either the source, the link between source and target, or the target were not able to keep up with the rate of demand. A 1720 error is a heartbeat or cluster partnership communication error. This error tends to be more serious, as failing communication between your cluster partners involves some extended diagnostic time.

9.4.1 1920 error

Let us focus first on the 1920 error. A 1920 error (event ID 050010) can have several triggers. Official probable cause projections are:

- Primary 2145 cluster or SAN fabric problem (10%)

- Primary 2145 cluster or SAN fabric configuration (10%)
- Secondary 2145 cluster or SAN fabric problem (15%)
- Secondary 2145 cluster or SAN fabric configuration (25%)
- Inter-cluster link problem (15%)
- Inter-cluster link configuration (25%)

In practice, the error that is most often overlooked is latency. Global Mirror has a round-trip-time tolerance limit of 80 ms. A message sent from your source SAN Volume Controller cluster to your target SAN Volume Controller Cluster and the accompanying acknowledgement must have a total time of 80 ms or 40 ms each way (for Version 4.1.1.x and later.)

Round-trip time: For Version 4.1.0.x and earlier, this limit was 68 ms or 34 ms one way for Fibre Channel extenders, and for SAN routers it was 10 ms one way or 20 ms round trip. Make sure to use the correct values for the correct versions!

The primary component of your round-trip time is the physical distance between sites. For every 1000 km (621.36 miles), there is a 5 ms delay. This delay does not include the time added by equipment in the path. Every device adds a varying amount of time, depending on the device, but expect about 25 µs for pure hardware devices. For software-based functions (such as compression implemented in software), the delay added tends to be much higher (usually in the millisecond plus range.)

Consider an example. Company A has a production site that is 1900 km distant from their recovery site. Their network service provider uses a total of five devices to connect the two sites. In addition to those devices, Company A employs a SAN Fibre Channel Router at each site to provide FCIP to encapsulate the Fibre Channel traffic between sites. There are now seven devices, and 1900 km of distance delay. All the devices add 200 µs of delay *each way.* The distance adds 9.5 ms each way, for a total of 19 ms. Combined with the device latency, that is 19.4 ms of *physical* latency at a minimum. This latency is under the 80 ms limit of Global Mirror, but this number is the best case number. Link quality and bandwidth play a significant role here. Your network provider likely guarantees a latency maximum on your network link; be sure to stay below the Global Mirror RTT limit. You can easily double or triple the expected physical latency with a lower quality or lower bandwidth network link. As a result you are suddenly within range of exceeding the limit the moment a large flood of I/O happens that exceeds the bandwidth capacity you have in place.

When you get a 1920 error, always check the latency first. Keep in mind that the FCIP routing layer can introduce latency if it is not properly configured. If your network provider reports a much lower latency, this report could be an indication of a problem at your FCIP Routing layer. Most FCIP Routing devices have built-in tools to allow you to check the RTT. When checking latency, remember that TCP/IP routing devices (including FCIP routers) report RTT using standard 64-byte ping packets.

Figure 9-68 on page 398 shows why the effective transit time should only be measured using packets large enough to hold a Fibre Channel frame. This packet size is 2148 bytes (2112 bytes of payload and 36 bytes of header) and you should allow some additional capacity to be safe, as different switching vendors have optional features that might increase this size. After you have verified your latency using the correct packet size, proceed with normal hardware troubleshooting.

Packet Size	Link Size	Serialization Delay (Time Required to Send Data)	Unit
64	256 Kbps	2.0E+03	microseconds
64	1.5 Mbps	3.4E+02	microseconds
64	100 Mbps	5.1E+00	microseconds
64	155 Mbps	3.3E+00	microseconds
64	622 Mbps	8.2E-01	microseconds
64	1 Gbps	5.1E-04	microseconds
64	10 Gbps	5.1E-05	microseconds
1500	256 Kbps	4.7E+04	microseconds
1500	1.5 Mbps	8.0E+03	microseconds
1500	100 Mbps	1.2E+02	microseconds
1500	155 Mbps	7.7E+01	microseconds
1500	622 Mbps	1.9E+01	microseconds
1500	1 Gbps	1.2E+01	microseconds
1500	10 Gbps	1.2E+00	microseconds
2148	256 Kbps	6.7E+04	microseconds
2148	1.5 Mbps	1.1E+04	microseconds
2148	100 Mbps	1.7E+02	microseconds
2148	155 Mbps	1.1E+02	microseconds
2148	622 Mbps	2.8E+01	microseconds
2148	1 Gbps	1.7E+01	microseconds
2148	10 Gbps	1.7E-03	microseconds

Figure 9-68 The effect of packet size (in bytes) versus the link size

Before you proceed, take a quick look at the second largest component of your round-trip-time, that is, serialization delay. Serialization delay is the amount of time required to move a packet of data of a specific size across a network link of a given bandwidth. This delay is based on a simple concept: the time required to move a specific amount of data decreases as the data transmission rate increases.

In Figure 9-68, there are orders of magnitude of difference between the different link bandwidths. It is easy to see how 1920 errors can arise when your bandwidth is insufficient and why you should never use a TCP/IP `ping` to measure RTT for FCIP traffic.

Figure 9-68 compares the amount of time in microseconds required to transmit a packet across network links of varying bandwidth capacity. Three packet sizes are used:

- 64 bytes: The size of the common ping packet
- 1500 bytes: The size of the standard TCP/IP packet
- 2148 bytes: The size of a Fibre Channel frame

Remember that your path MTU affects the delay incurred in getting a packet from one location to another, when it causes fragmentation or is too large and causes too many retransmits when a packet is lost.

9.4.2 1720 error

The 1720 error (event ID 050020) is the other primary error code of Remote Copy. Because the term “System Partnership” implies that all involved virtualization systems are partners, they must be able to communicate with each other. When a partner on either side stops communicating, you see a 1720 error appear in your error log. According to official documentation, there are no likely field replaceable unit breakages or other causes.

In practice, the source of this error is most often a fabric problem or a problem the network path between your partners. When you receive this error, if your fabric has more than 64 HBA ports zoned, you should check your fabric configuration for zoning of more than one HBA port

for each node per I/O group. One port for each node per I/O group associated with the host is the recommended zoning configuration for fabrics. For those fabrics with 64 or more host ports, this recommendation becomes a rule. You must follow this zoning rule or the configuration is technically unsupported.

Improper zoning leads to SAN congestion, which can inhibit remote link communication intermittently. Checking the zero buffer credit timer through IBM Tivoli Storage Productivity Center and comparing its value against your sample interval might reveal potential SAN congestion. Anytime a zero buffer credit timer is above 2% of the total time of the sample interval, it is likely to cause problems.

Next, always ask your network provider to check the status of the link. If the link is okay, watch for repetition of this error. It is possible in a normal and functional network setup to have occasional 1720 errors, but multiple occurrences point to a larger problem.

If you receive multiple 1720 errors, recheck your network connection and then check the IBM Flex System V7000 Storage Node partnership information to verify their status and settings. Perform diagnostic tests for every piece of equipment in the path between the two systems. It often helps to have a diagram showing the path of your replication from both logical and physical configuration viewpoints.

If your investigation fails to resolve your Remote Copy problems, you should contact your IBM support representative for a complete analysis.

9.5 Managing Remote Copy using the GUI

The IBM Storwize V7000 storage system provides a separate function icon for copy service management. There are two windows for managing Remote Copy, which are accessed through the Copy Services function icon:

- Remote Copy
- Partnerships

As the name implies, these two windows are used to manage Remote Copy and the partnership for the copy relationship.

9.5.1 Managing cluster partnerships

The Partnership window is used to manage a partnership between the mirroring clusters. To access the Partnership window, click the **Copy Services** function icon and click **Partnerships** (Figure 9-69 on page 400).

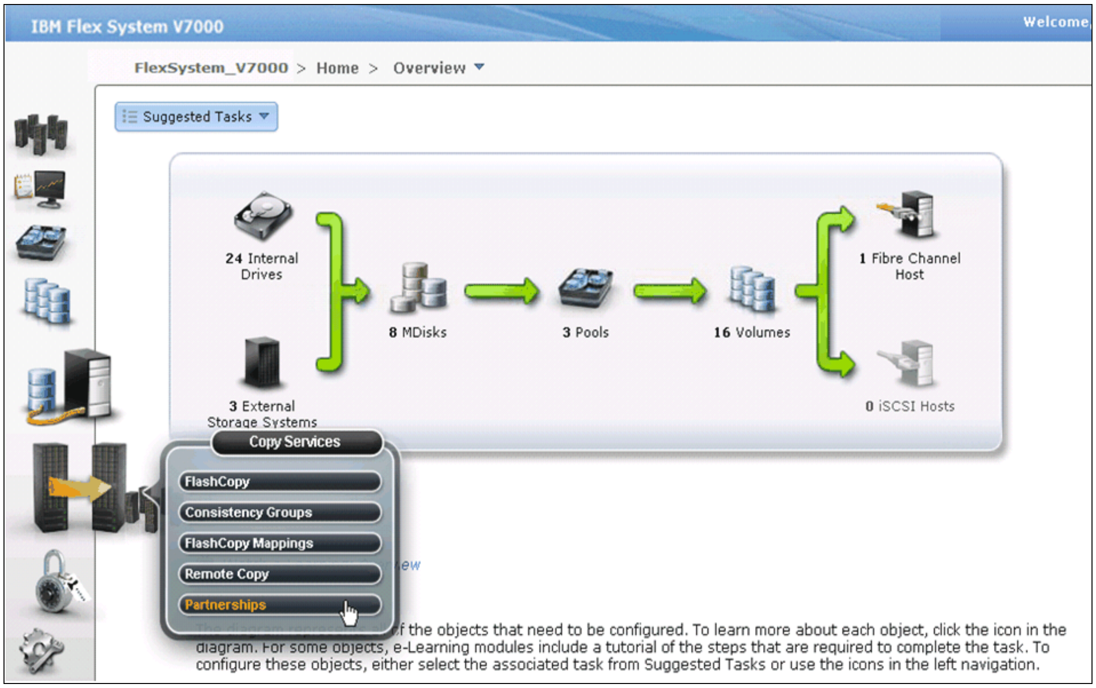


Figure 9-69 Partnership window

Creating a partnership

No partnership is defined in our example (Figure 9-70), so you must create a partnership between the IBM Flex System V7000 Storage Node system and the other mirroring member. Click **New Partnership** in the Partnership window.

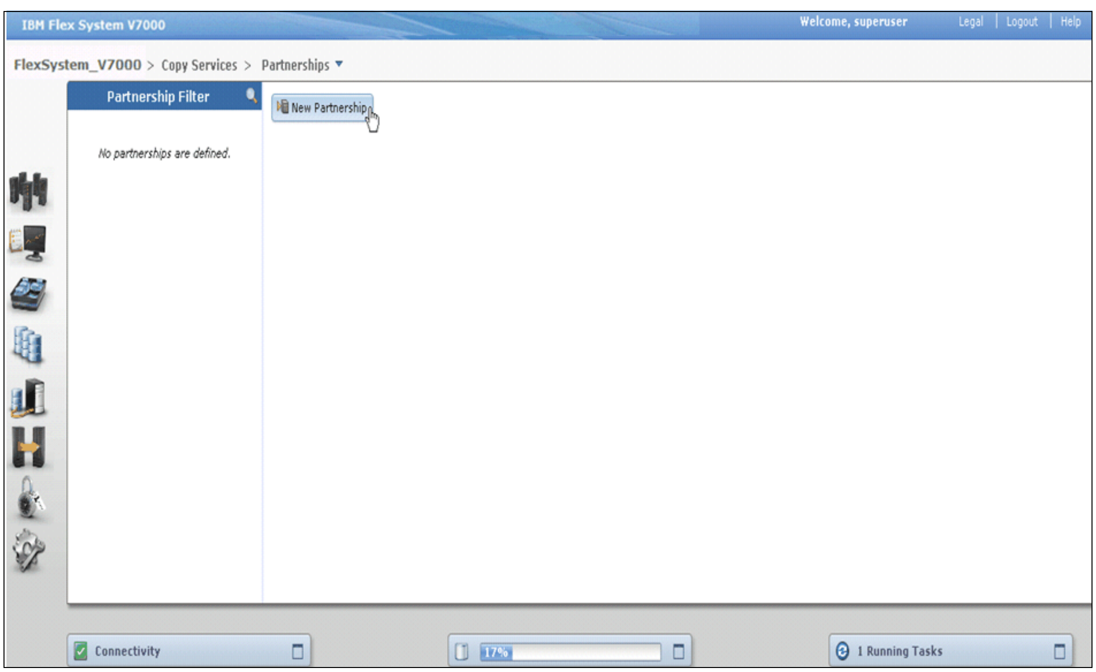


Figure 9-70 Create a cluster partnership

If there is no partnership candidate, an error window opens (Figure 9-71 on page 401).

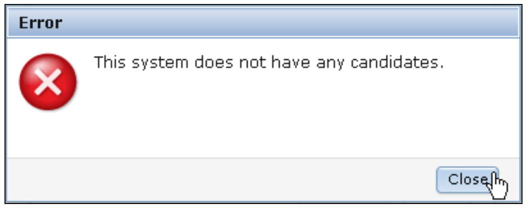


Figure 9-71 No candidates are available to create a partnership

Check the zoning and the system status and make sure that the clusters can "see" each other. Then you can create your partnership by selecting the appropriate remote storage system (Figure 9-72), and defining the available bandwidth between both systems.

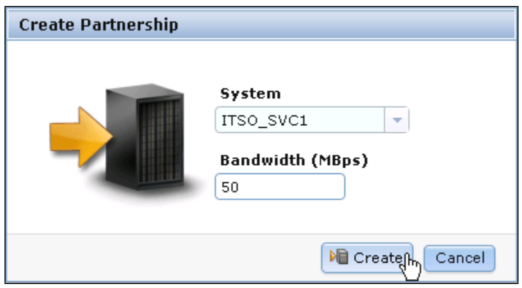


Figure 9-72 Select a remote storage system for a new partnership

In our example we have built the relationship between FlexSystem_V7000 and ITSO_SVC1 The bandwidth you need to input here is used by the background copy process between the clusters in the partnership. To set the background copy bandwidth optimally, make sure that you consider all three resources (the primary storage, the inter-cluster link bandwidth, and the auxiliary storage) to avoid overloading them and affecting the foreground I/O latency.

Click **Create** and the partnership definition is complete on the IBM Flex System V7000 Storage Node. You can find the partnership listed in the left pane of the Partnership window, and if you select the partnership, more information for this partnership is displayed on the right (Figure 9-73 on page 401).

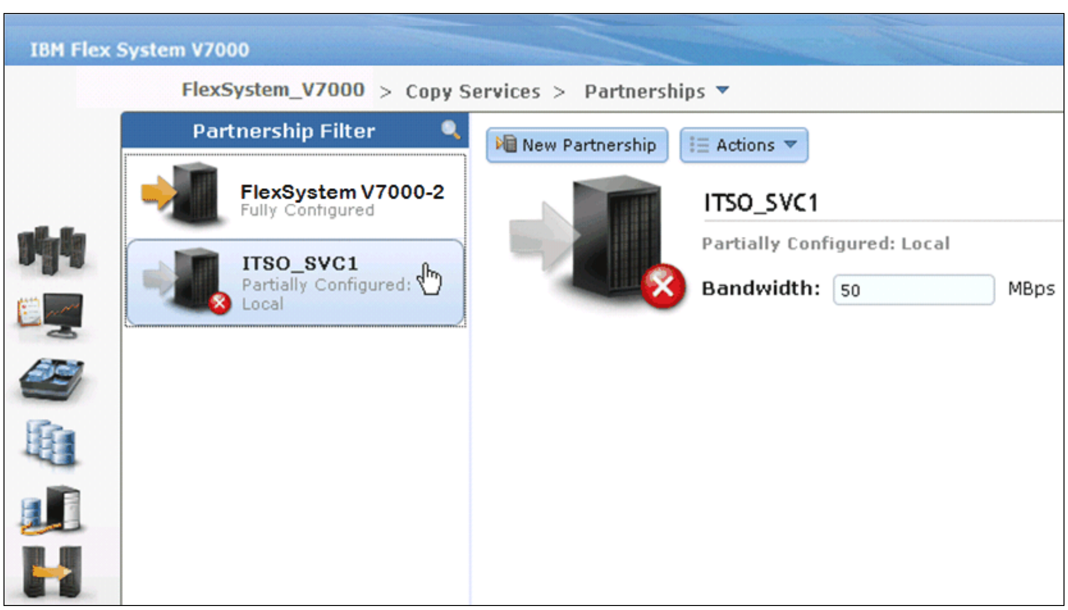


Figure 9-73 Partially configured partnership

Complete the same steps on the second storage system that becomes a fully configured partner.

The Remote Copy partnership is now implemented between two IBM Flex System V7000 Storage Node systems and both systems are ready for further configuration of Remote Copy relationships (Figure 9-74).

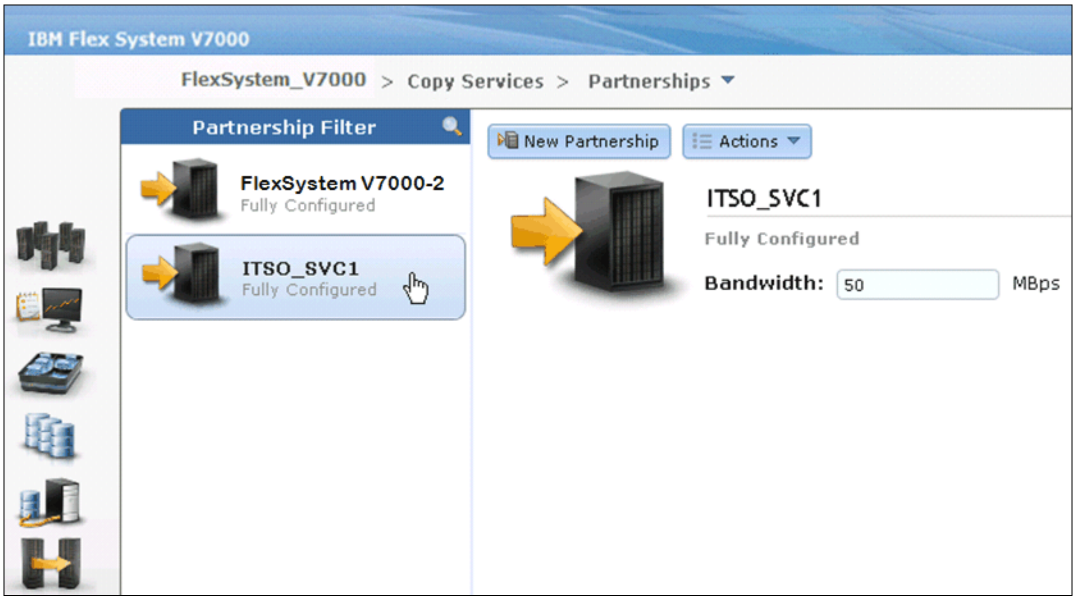


Figure 9-74 Fully configured partnership

You can also change the bandwidth setting for the partnership in the Partnerships window (Figure 9-75 on page 402). Click **Apply Changes** to confirm your modification.

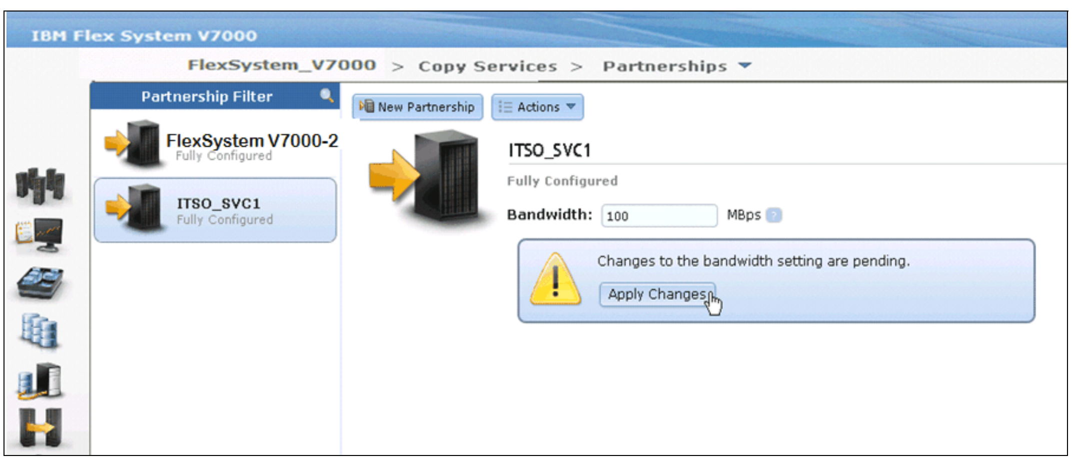


Figure 9-75 Apply changes to the partnership bandwidth

Stopping and starting a partnership

You can stop the partnership by clicking **Stop Partnership** from the **Actions** drop-down menu (Figure 9-76). If you stop the partnership, the relationship using this partnership is disconnected.

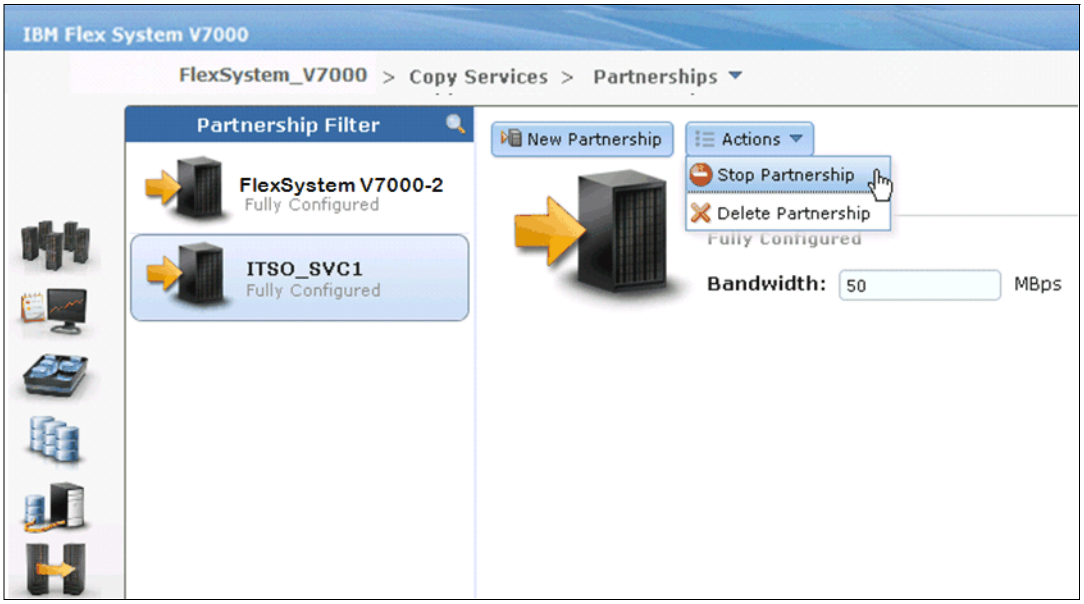


Figure 9-76 Stop the partnership

After you have stopped the partnership, your partnership is listed as Fully Configured: Stopped (Figure 9-77).

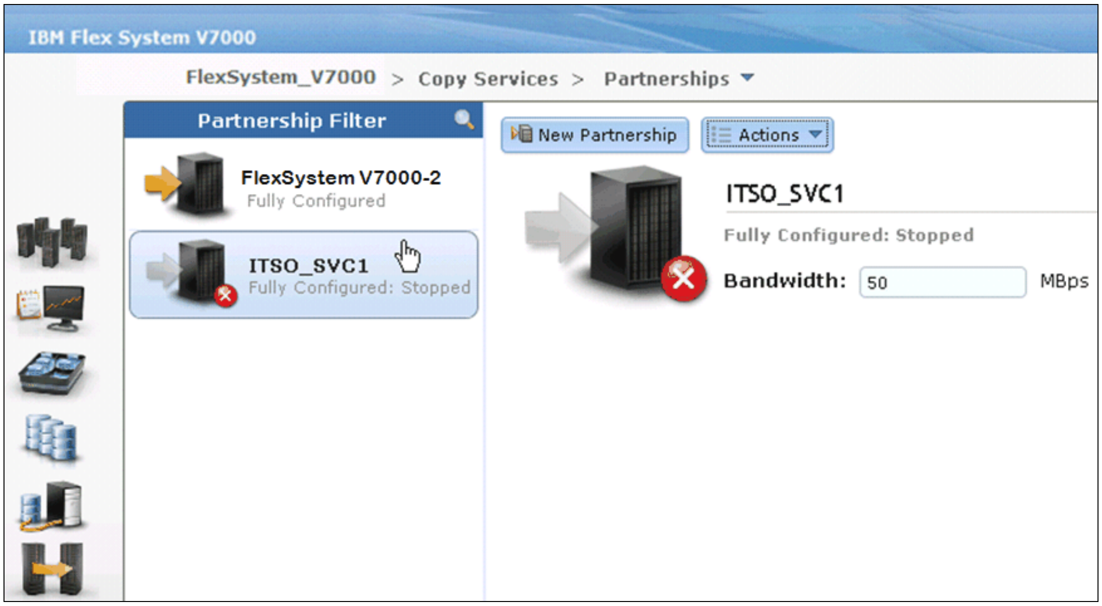


Figure 9-77 Fully configured partnership in stopped state

You can restart a stopped partnership by clicking **Start Partnership** from the **Actions** drop-down menu (Figure 9-78).

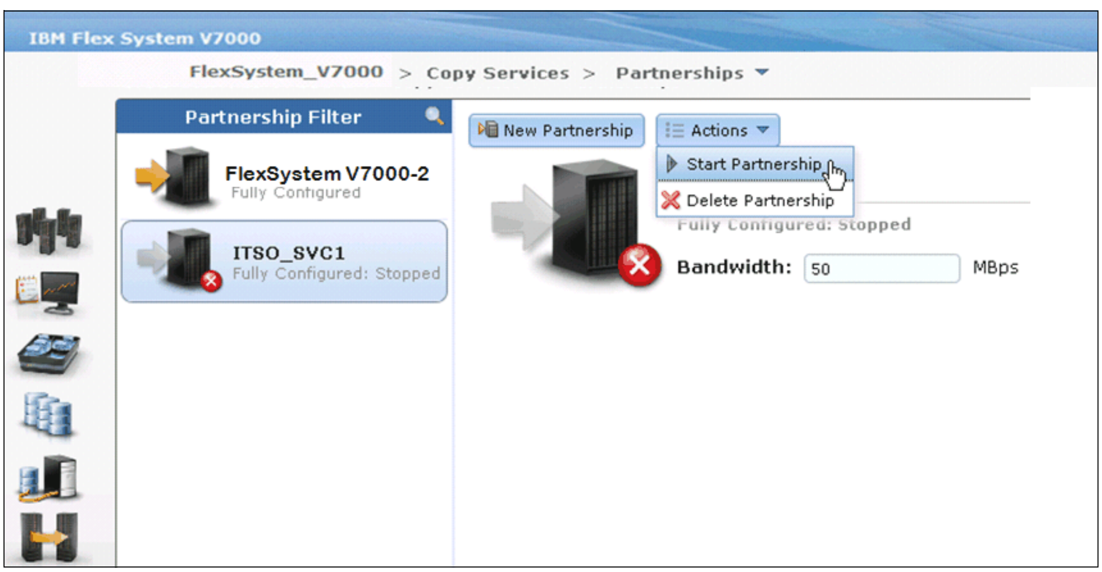


Figure 9-78 Start partnership

The partnership returns to the fully configured status when it has been restarted.

9.5.2 Deleting a partnership

You can delete a partnership by selecting **Delete Partnership** from the **Actions** drop-down menu (Figure 9-79).

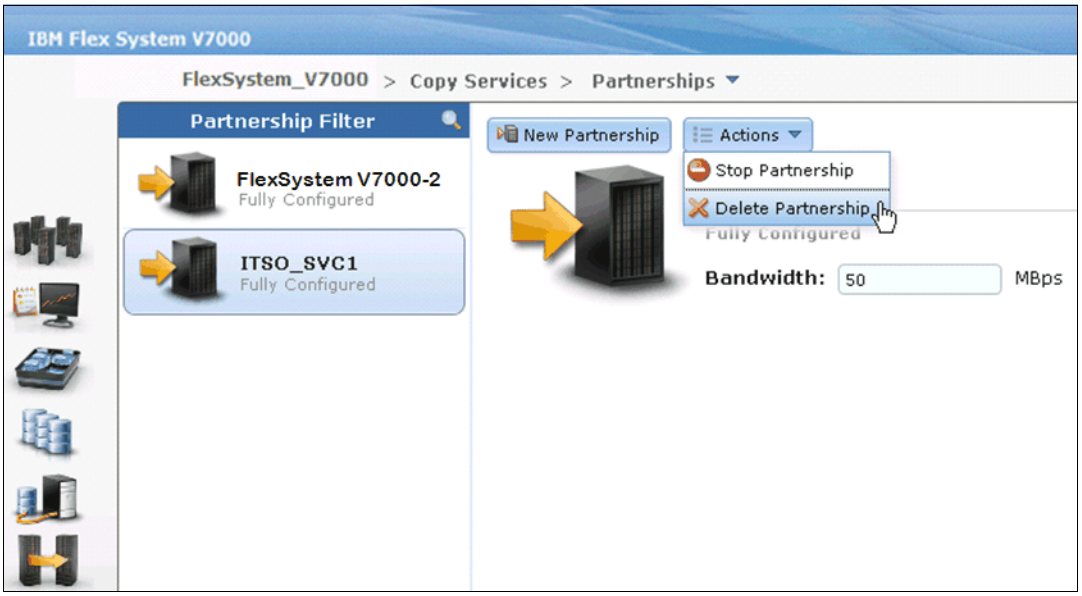


Figure 9-79 Delete a partnership

Managing stand-alone Remote Copy relationships

A Remote Copy relationship can be defined between two volumes, where one is the master (source) and the other one is the auxiliary (target) volume. Usage of Remote Copy auxiliary volumes as Remote Copy master volumes is not allowed.

Open the Remote Copy window to manage Remote Copy by clicking the **Copy Services** icon and clicking **Remote Copy** (Figure 9-80).

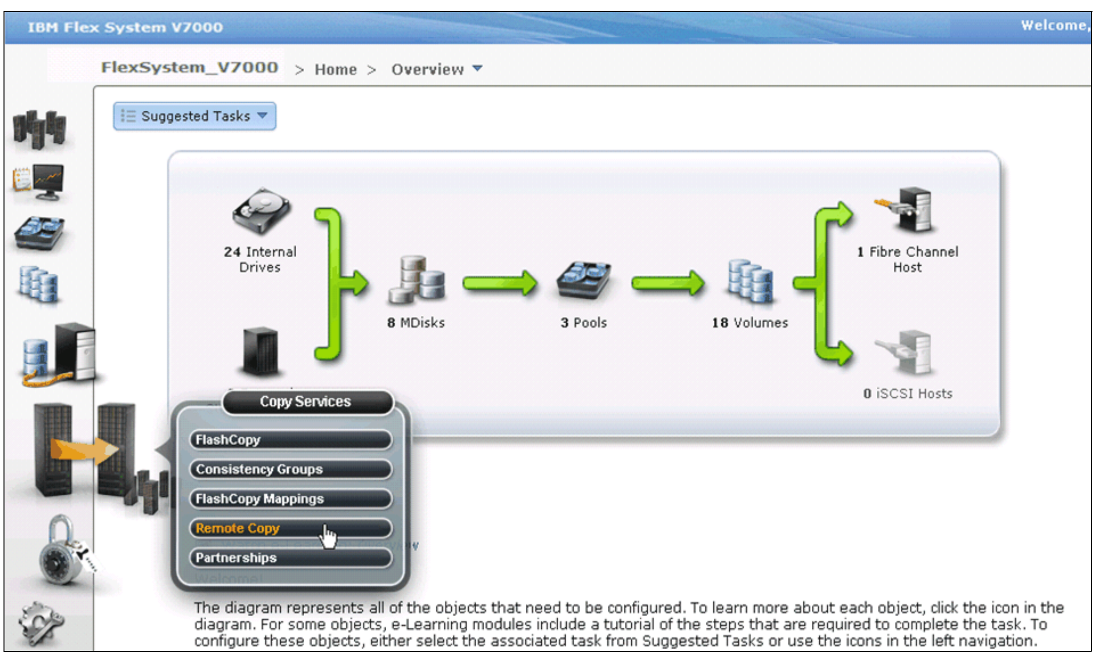


Figure 9-80 Open Remote Copy window

The Remote Copy window (Figure 9-81) is where you can manage Remote Copy relationships and Remote Copy consistency groups.

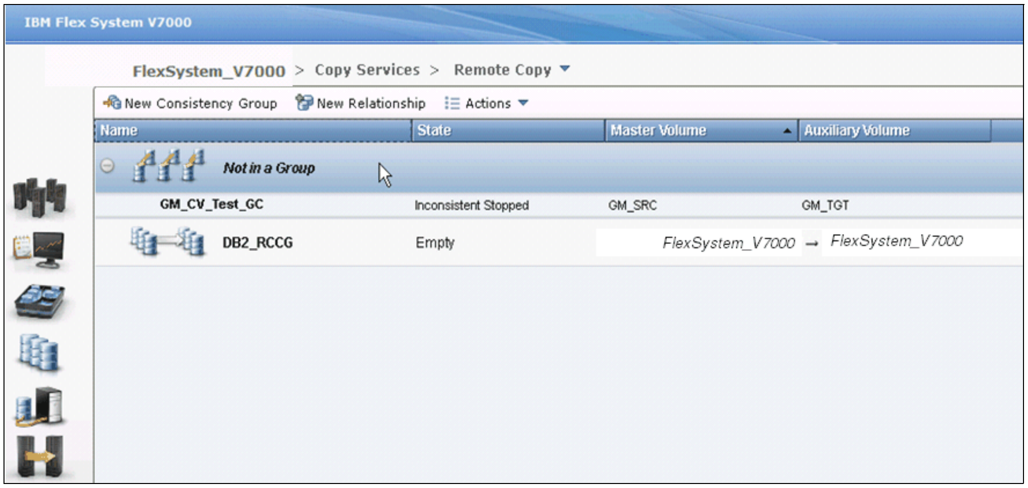


Figure 9-81 Remote Copy window

In the left pane of the Remote Copy window, there is a consistency group filter to list Remote Copy consistency groups that meet your requirements. In the right pane of the Remote Copy window, you can discover the properties of a consistency group and the Remote Copy relationships in it. You can also take actions on the Remote Copy relationships and Remote Copy consistency group. Click **Not in a Group**, and all the Remote Copy relationships that are not in any Remote Copy consistency groups are displayed on the right.

Creating stand-alone Remote Copy relationships

To create a Remote Copy relationship, click **New Relationship** at the top of the Remote Copy window (Figure 9-81). A wizard opens and guides you through the Remote Copy relationship creation process.

As shown in Figure 9-82, you need to set the Remote Copy relationship type first. Based on your requirements, you can select Metro Mirror (synchronous replication) or Global Mirror (asynchronous replication). Select the appropriate replication type and click **Next**.

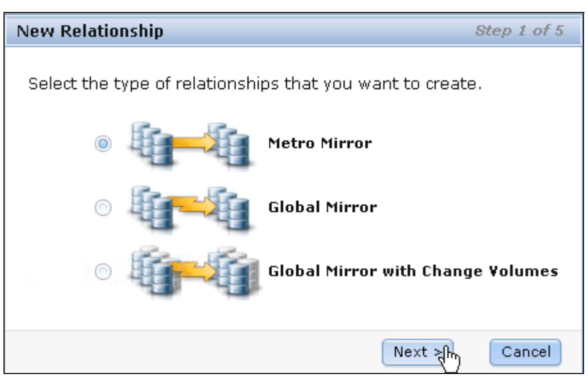


Figure 9-82 Select the appropriate Remote Copy type

You need to select your Remote Copy auxiliary (target) storage system, the local system, or the already defined second storage system as the Remote Copy partner.

Note: The use of creating a mirror relationship on the local system is frequently used to develop an understanding of the remote copy procedures and usage, as well as for testing purposes.

In our example (Figure 9-83), choose another system to build an inter-cluster relationship. Click **Next** to continue.

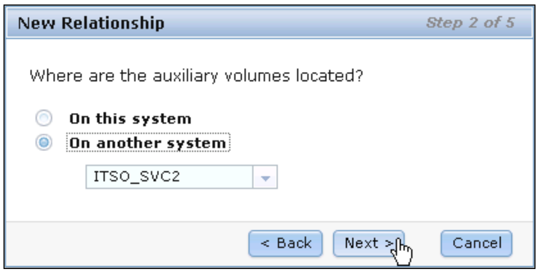


Figure 9-83 Select Remote Copy partner

The Remote Copy master and auxiliary volume need to be specified. Both volumes must have the same size. As shown in Figure 9-84, the system offers only appropriate auxiliary candidates with the same volume size as the selected master volume. After you select the volumes based on your requirement, click **Add**.

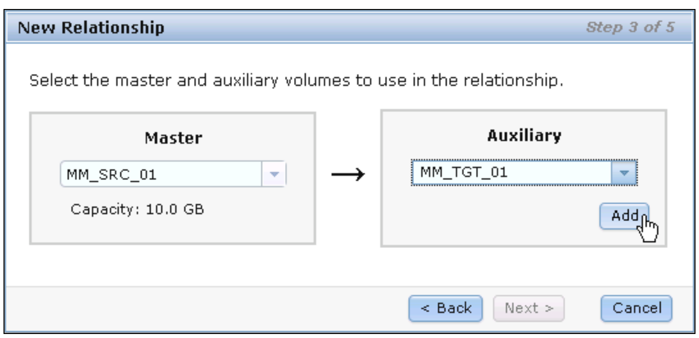


Figure 9-84 Select the master and auxiliary volume

You can define multiple and independent relationships by clicking **Add**. You can remove a relationship by clicking the red cross. In our example, create two independent Remote Copy relationships (Figure 9-85).

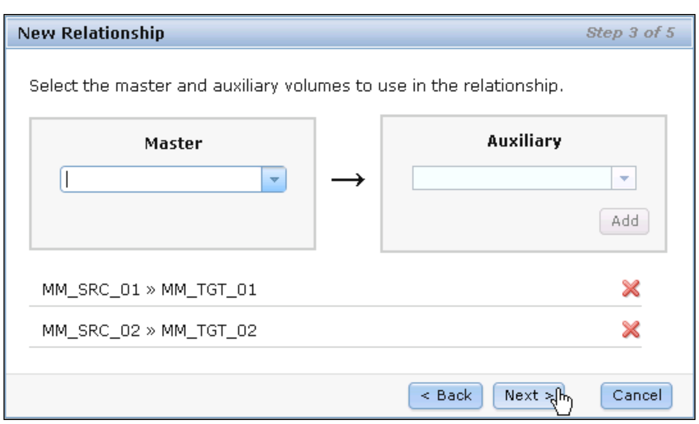


Figure 9-85 Define multiple independent relationships

A window opens and asks if the volumes in the relationship are already synchronized. In most situations, the data on the master volume and on the auxiliary volume are not identical, so click **No** and click **Next** to enable an initial copy (Figure 9-86).

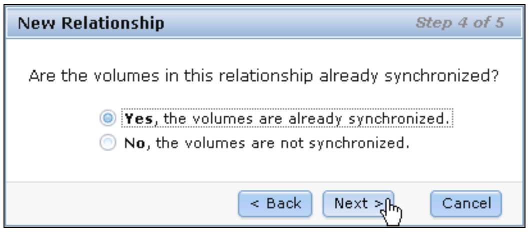


Figure 9-86 Activate initial data copy

If you select **Yes, the volumes are already synchronized** in this step, a warning message opens (Figure 9-87). Make sure that the volumes are truly identical, and then click **OK** to continue.

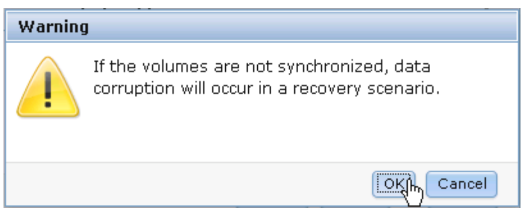


Figure 9-87 Warning message to make sure that the volumes are synchronized

You can choose to start the initial copying progress now, or wait to start it at a later time. In our example, select **Yes, start copying now** and click **Finish** (Figure 9-88).

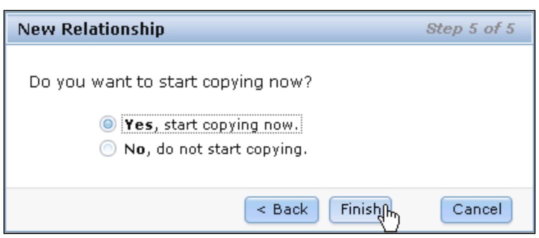


Figure 9-88 Choose if you want to start copying now or later

After the Remote Copy relationships creation completes, two independent Remote Copy relationships are defined and displayed in the Not in a Group list (Figure 9-89).

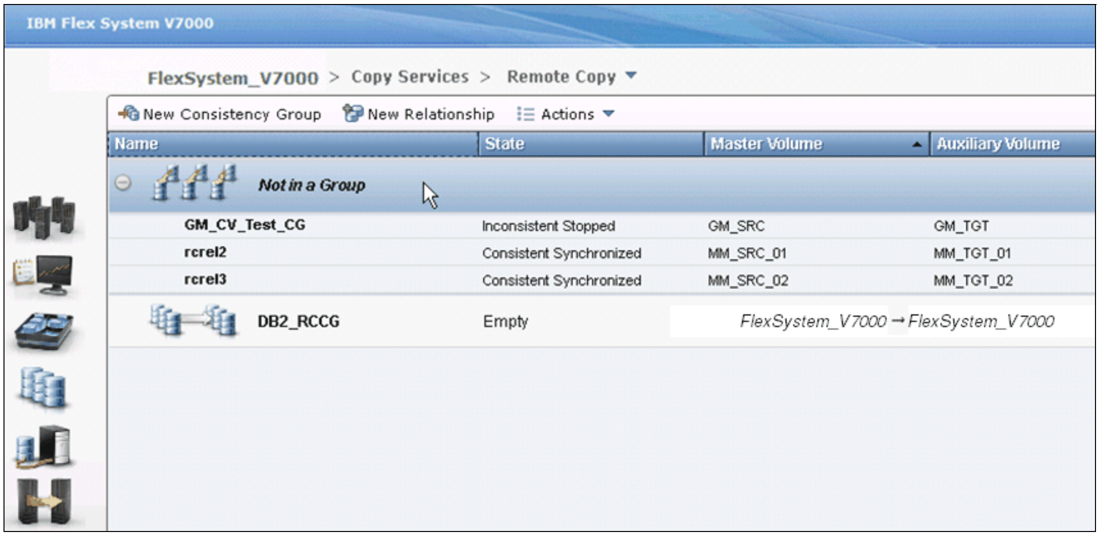


Figure 9-89 Remote Copy relationship create completes

Optionally, you can monitor the ongoing initial synchronization in the Running Tasks status indicator (Figure 9-90).

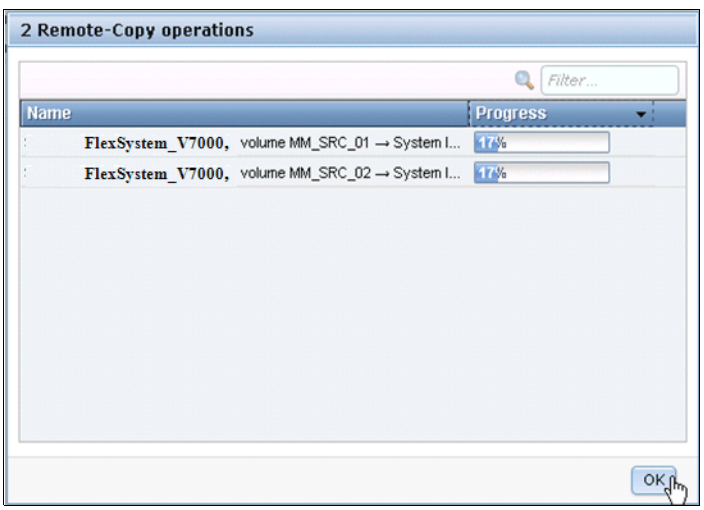


Figure 9-90 Remote copy initialization progress through Running Tasks

Stopping a stand-alone Remote Copy relationship

The Remote Copy relationship can be stopped by selecting the relationship and clicking **Stop** from the **Actions** drop-down menu (Figure 9-91 on page 410).

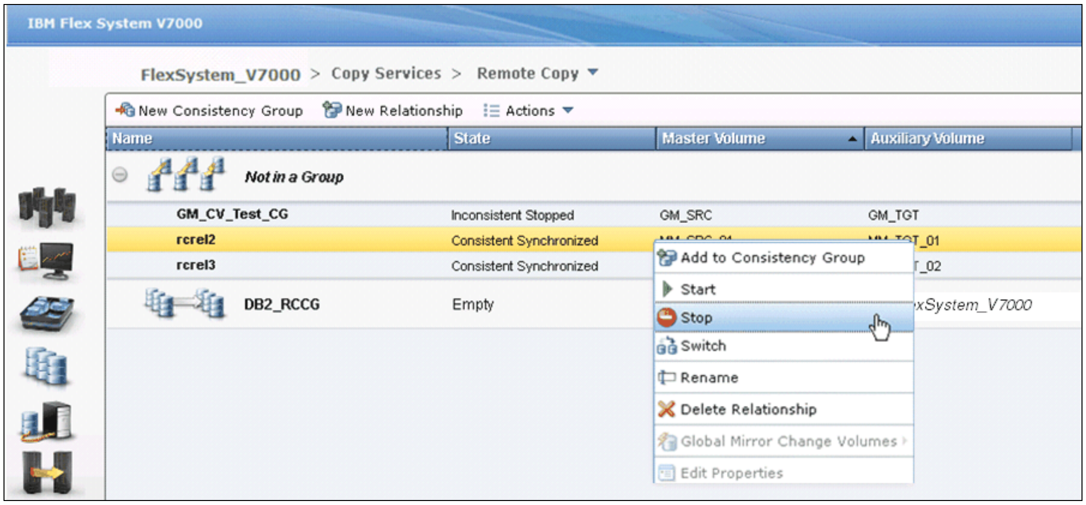


Figure 9-91 Stop Remote Copy relationship

Allow secondary read/write access, if required, and click **Stop Relationship** (Figure 9-92).

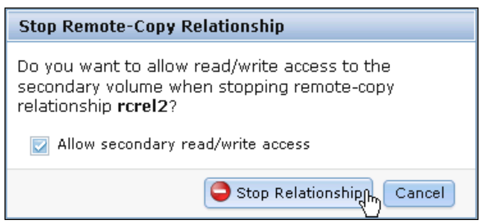


Figure 9-92 Option to allow secondary read/write access

After the stop completes, the state of the Remote Copy relationship is changed from Consistent Synchronized to Idling (Figure 9-93). Read / write access to both volumes is now allowed.

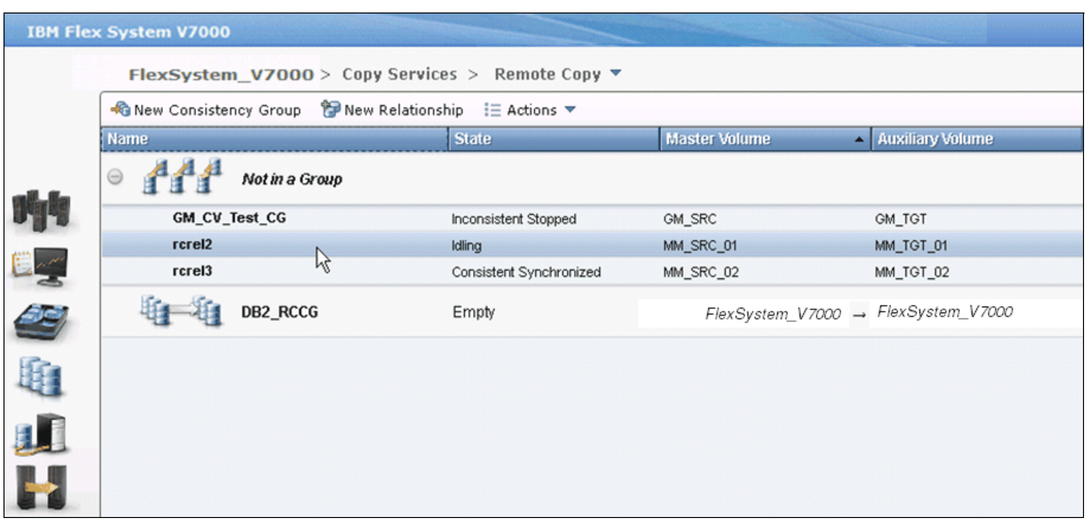


Figure 9-93 Remote Copy relationship stop completes

Starting a stand-alone Remote Copy relationship

The Remote Copy relationship can be started by selecting the relationship and clicking **Stop** from the **Actions** drop-down menu (Figure 9-94).

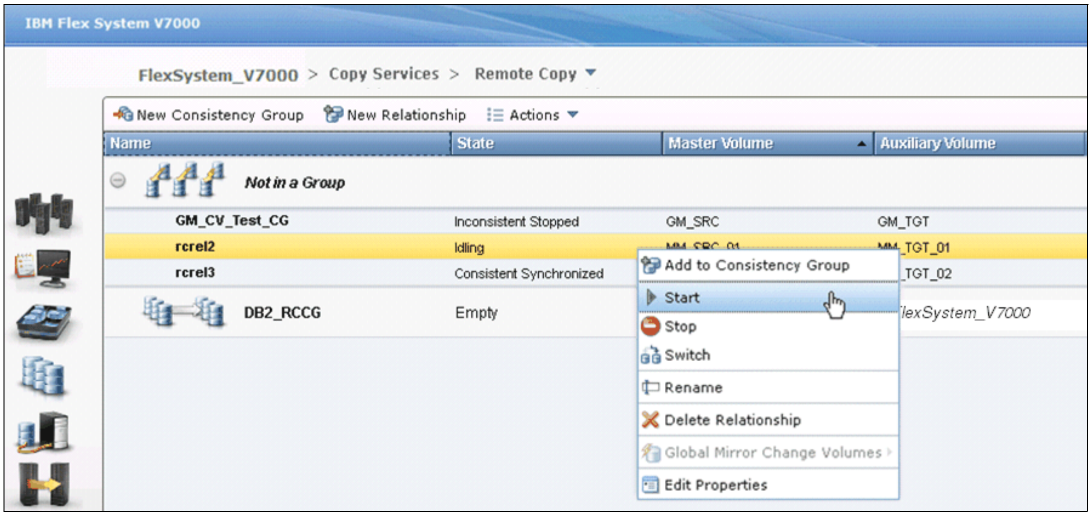


Figure 9-94 Start a Remote Copy relationship

When starting a Remote Copy relationship, the most important item is selecting the copy direction. Both master and auxiliary volumes could be the primary. Make your decision based on your requirements and click **Start Relationship**. In our example, choose the master volume to be the primary (Figure 9-95).

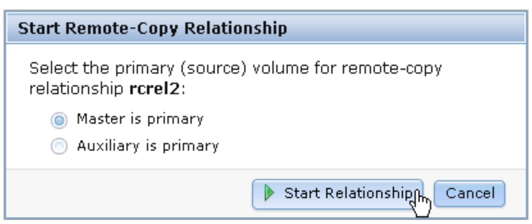


Figure 9-95 Choose the copy direction

Switching the direction of a stand-alone Remote Copy relationship

The copy direction of the Remote Copy relationship could be switched by selecting the relationship and clicking **Switch** from the **Actions** drop-down menu (Figure 9-96).

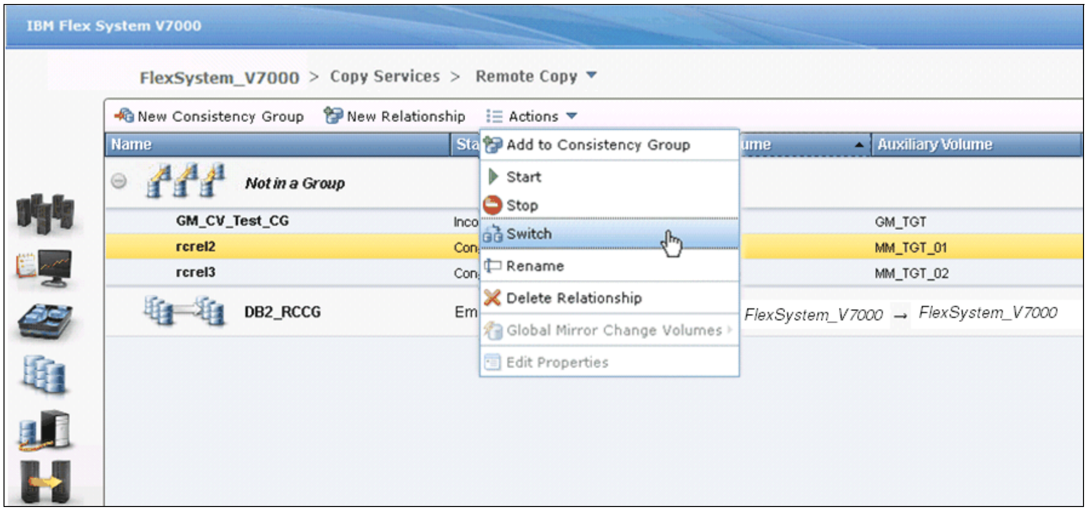


Figure 9-96 Switch Remote Copy relationship

A warning message opens and shows you the consequences of this action (Figure 9-97). If you switch the Remote Copy relationship, the copy direction of the relationship becomes the opposite, that is, the current primary volume becomes the secondary, while the current secondary volume becomes the primary. Write access to the current primary volume is lost and write access to the current secondary volume is enabled. If it is not a disaster recovery situation, you need to stop your host I/O to the current primary volume in advance. Make sure that you are prepared for the consequences, and if so, click **OK** to continue.

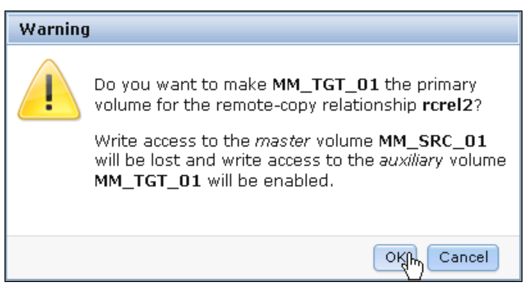


Figure 9-97 Warning message for switching direction of a Remote Copy relationship

After the switch completes, your Remote Copy relationship is tagged (Figure 9-98 on page 413), and shows you that the primary volume in this relationship has been changed.

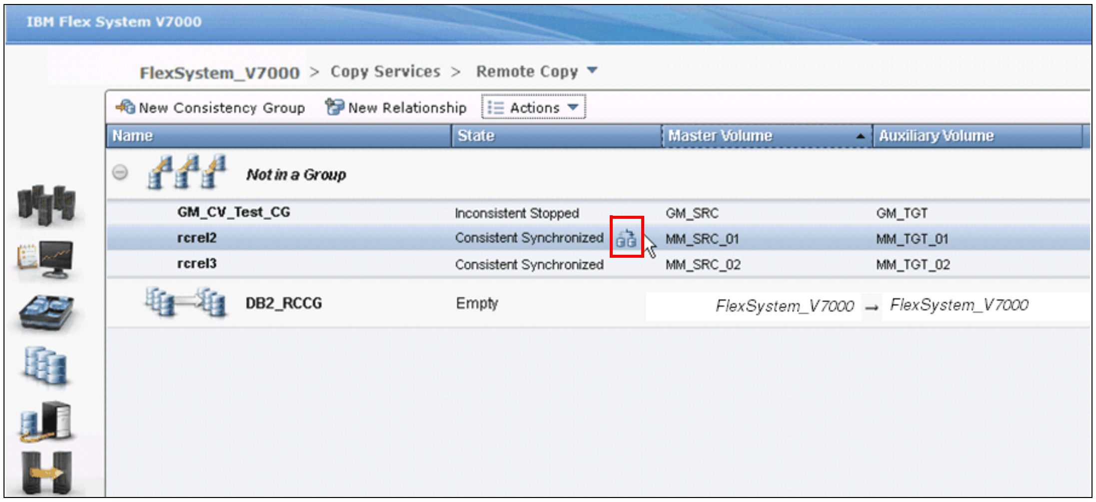


Figure 9-98 Note the switch icon on the state of the relationship

Renaming a stand-alone Remote Copy relationship

The Remote Copy relationship can be renamed by selecting the relationship and clicking **Rename** from the **Actions** drop-down menu (Figure 9-99).

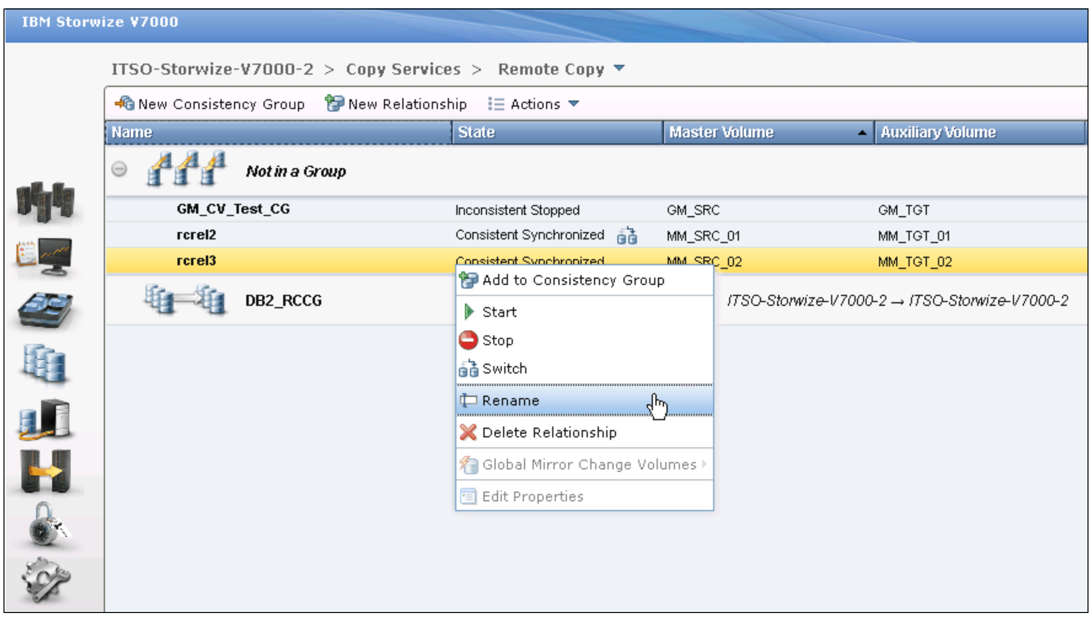


Figure 9-99 Rename the Remote Copy relationship

Enter the new name for the Remote Copy relationship and click **Rename** (Figure 9-100).

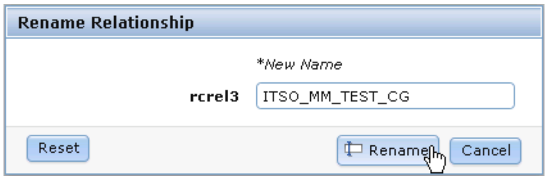


Figure 9-100 Enter a new name for the Remote Copy relationship

Deleting a stand-alone Remote Copy relationship

The Remote Copy relationship can be deleted by selecting the relationship and clicking **Delete Relationship** from the **Actions** drop-down menu (Figure 9-101).

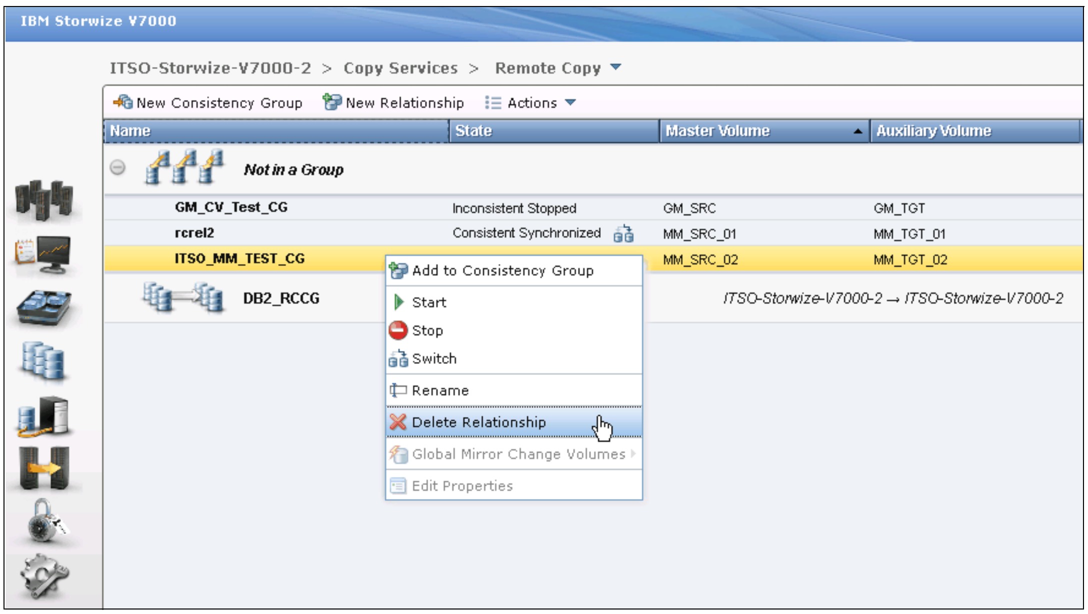


Figure 9-101 Delete a Remote Copy relationship

You must confirm this deletion by verifying the number of relationships to be deleted (Figure 9-102). Click **Delete** to proceed.

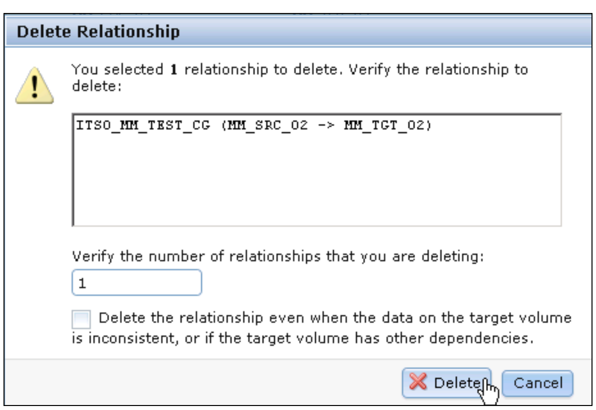


Figure 9-102 Confirm the relationship deletion

9.5.3 Managing a Remote Copy consistency group

A Remote Copy consistency group can be managed from the Remote Copy window as well.

Creating a Remote Copy consistency group

To create a Remote Copy consistency group, click **New Consistency Group** (Figure 9-103).

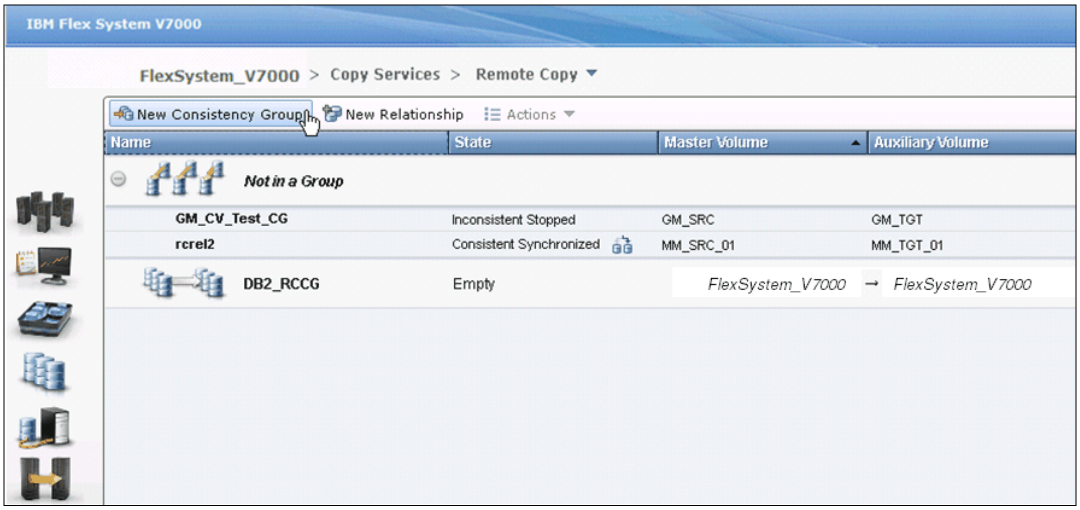


Figure 9-103 Create a consistency group

You need to enter a name for your new consistency group (Figure 9-104).

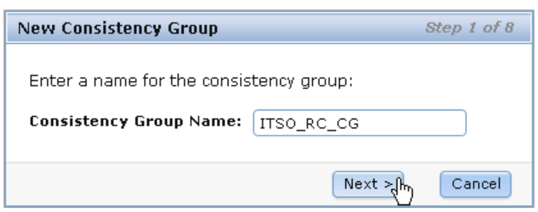


Figure 9-104 Enter a name for the new consistency group

You can choose to create an empty consistency group or, if you want, add Remote Copy relationships into the consistency group now (Figure 9-105). If you select **Yes, add relationships to this group**, you can select existing relationships or create new ones to add to the consistency group. In our example, create an empty consistency group, and add Remote Copy relationships to the consistency group afterward. Click **Finish** to proceed.

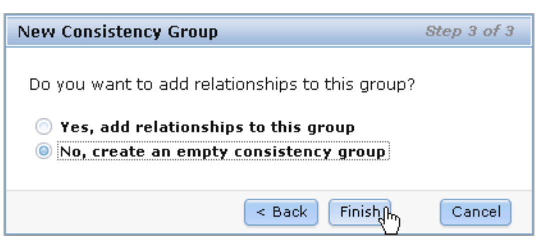


Figure 9-105 Create an empty consistency group

After the creation process completes, a new empty consistency group appears in the left pane of the Remote Copy window. Click the new consistency group, and you can find more information in the right pane (Figure 9-106).

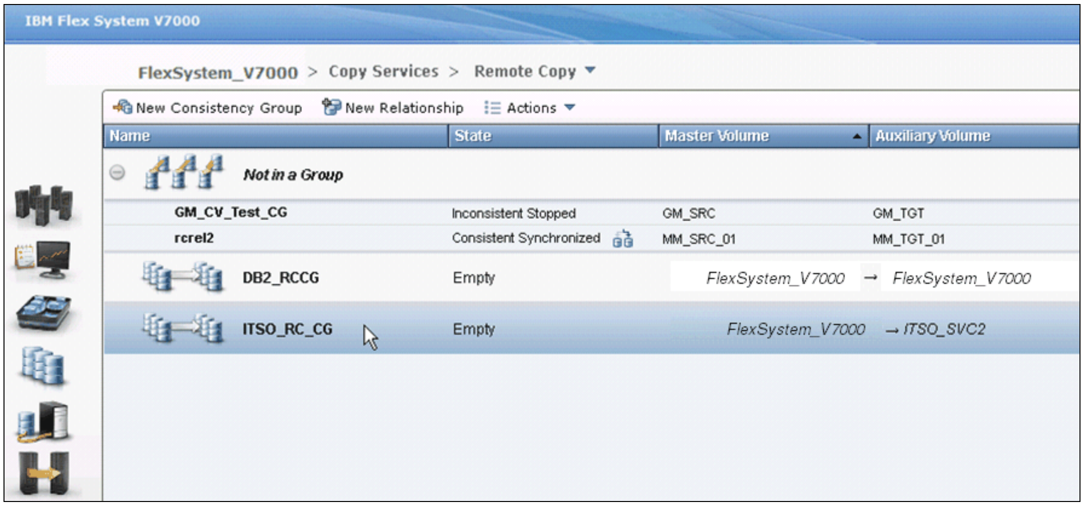


Figure 9-106 New Remote Copy consistency group

You can find the name and the status of the consistency group beside the Relationship function icon. It is easy to change the name of consistency group by clicking the name and entering a new one. At the upper right of the Remote Copy window, you can take actions on the Remote Copy consistency group. Below the Relationship function icon, you find all the Remote Copy relationships in this consistency group. The actions on the Remote Copy relationships can be applied here by using the **Actions** drop-down menu.

Adding Remote Copy to a consistency group

The Remote Copy relationships in the Not in a Group list could be added to a consistency group by selecting the volumes and clicking **Add to Consistency Group** from the **Actions** drop-down menu (Figure 9-107).

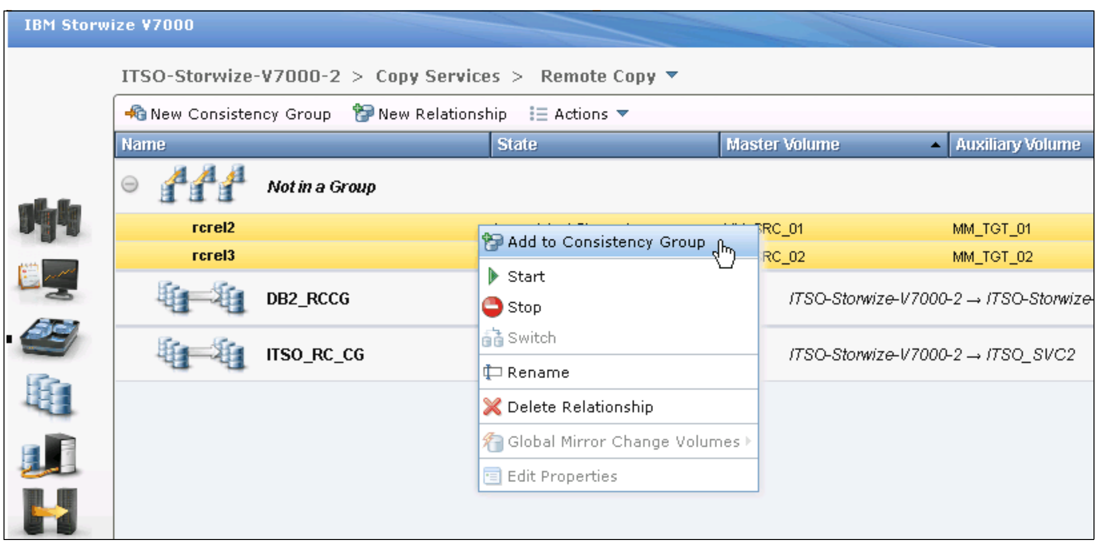


Figure 9-107 Add Remote Copy relationships to a consistency group

You must choose the consistency group to which to add the Remote Copy relationships. Based on your requirements, select the appropriate consistency group and click **Add to Consistency Group** (Figure 9-108).

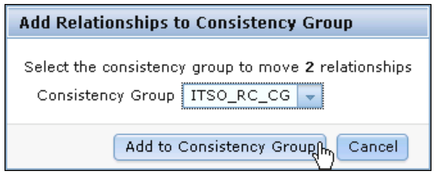


Figure 9-108 Choose the consistency group to add the remote copies to

Your Remote Copy relationships are now in the consistency group you selected.

Starting a consistency group

The Remote Copy relationship could be started by clicking **Start** from the **Actions** drop-down menu (Figure 9-109).

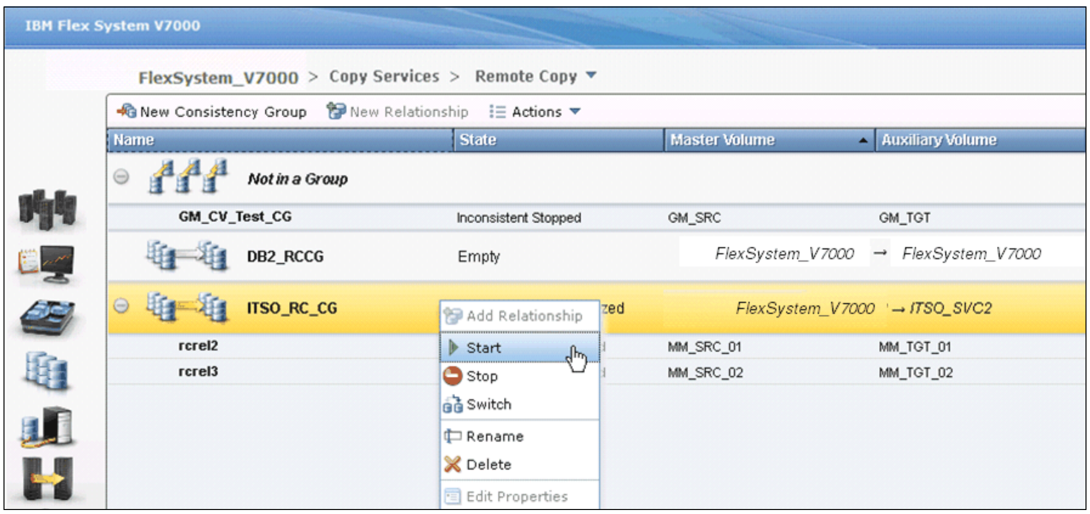


Figure 9-109 Start the consistency group

You can select the copy direction of the consistency group (as required) (Figure 9-110). Choose **Master is primary** and click **Start Consistency Group**. The consistency group starts copying data from the primary to the secondary.

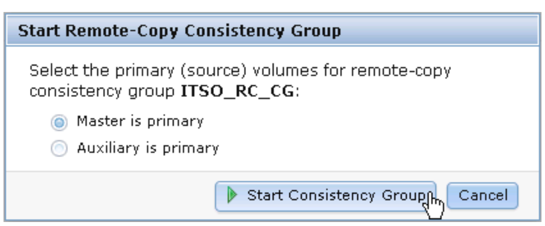


Figure 9-110 Choose the primary volumes to start the consistency group

Stopping a consistency group

The Remote Copy relationship can be stopped by clicking **Stop** in the **Actions** drop-down menu (Figure 9-111).

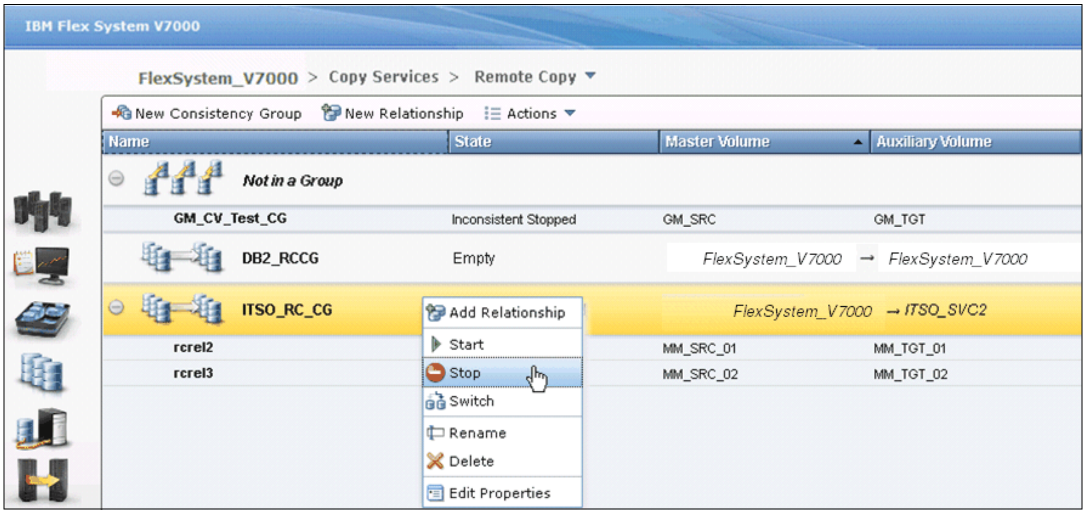


Figure 9-111 Stop the consistency group

You can allow read / write access to secondary volumes by selecting the check box (Figure 9-112) and clicking **Stop Consistency Group**.

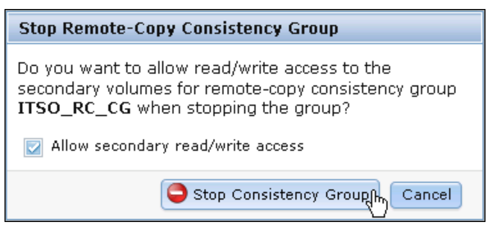


Figure 9-112 Confirm consistency group stop and decide to allow secondary read / write access

Switching a consistency group

As with the switch action on the Remote Copy relationship, you can switch the copy direction of the consistency group as well. To switch the copy direction of the consistency group, click **Switch** from the **Actions** drop-down menu (Figure 9-113).

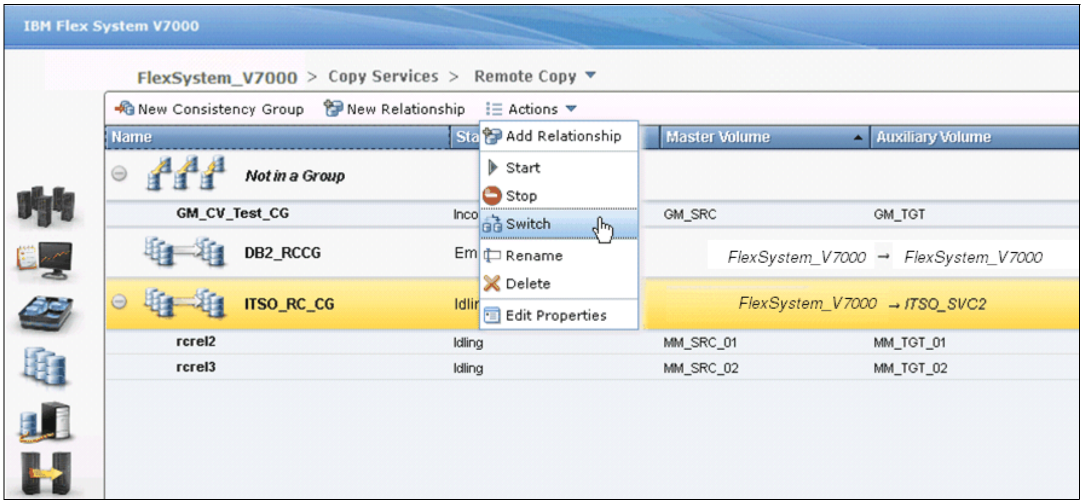


Figure 9-113 Switch the copy direction of a consistency group

A warning message opens (Figure 9-114). After the switch, the primary cluster in the consistency group is changed. Write access to current master volumes is lost, while write access to the current auxiliary volumes is enabled. Make sure that these settings are what you need, and if so, click **OK** to continue.

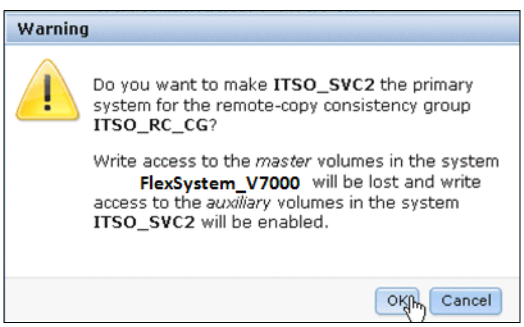


Figure 9-114 Warning message to confirm the switch

Removing Remote Copy relationships from a consistency group

The Remote Copy relationships can be removed from the consistency group by selecting the Remote Copy relationships and clicking **Remove from Consistency Group** from the **Actions** drop-down menu (Figure 9-115 on page 420).

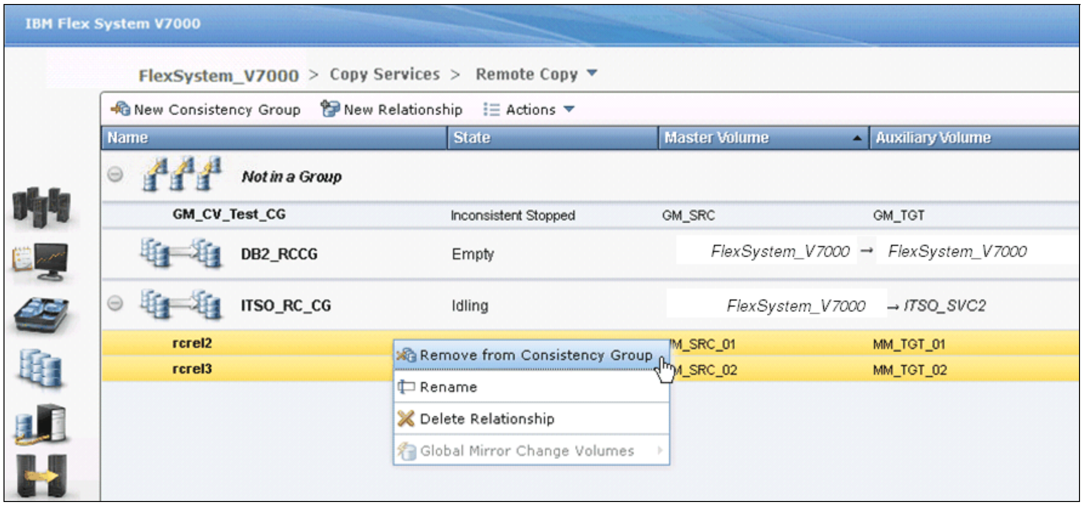


Figure 9-115 Remove Remote Copy relationships from a consistency group

You are prompted to confirm the Remote Copy relationships you want to delete from the consistency group (Figure 9-116). Make sure the Remote Copy relationships shown in the box are the ones you need to remove from the consistency group, and click **Remove** to proceed.

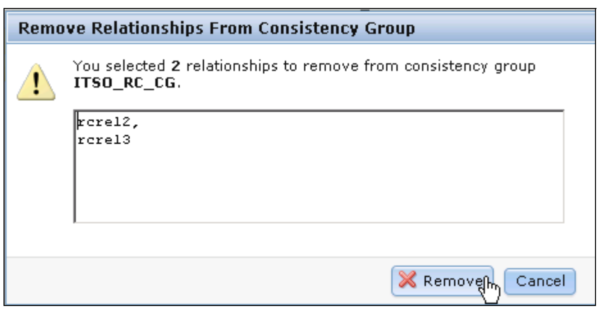


Figure 9-116 Confirm the relationships to remove from the Remote Copy consistency group

After the removal process completes, the Remote Copy relationships are deleted from the consistency group and displayed in the Not in a Group list.

Deleting a consistency group

The consistency group can be deleted by selecting **Delete** from the **Actions** drop-down menu (Figure 9-117 on page 421).

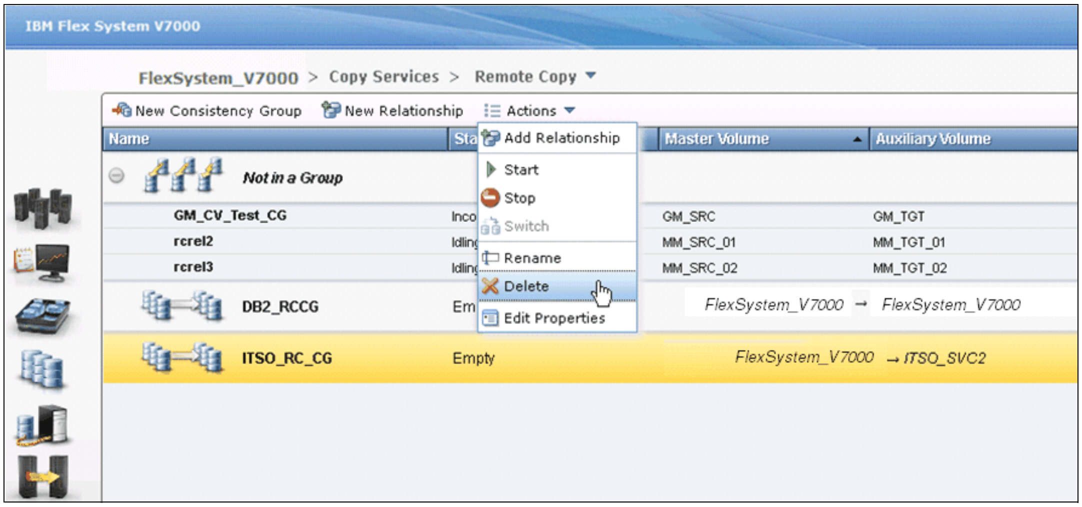


Figure 9-117 Delete a consistency group

You must confirm the deletion of the consistency group (Figure 9-118). Click **OK** if you are sure that this consistency group should be deleted.

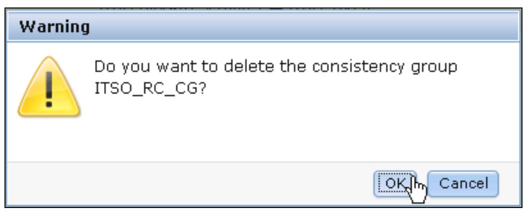


Figure 9-118 Warning to confirm deletion of the consistency group

The consistency group is deleted.

10

Volume Mirroring and Migration

This chapter discusses a number of the details and capabilities of the volume mirroring function. This feature was originally developed as a part of the IBM SAN Volume Controller (SVC) and is shared with the IBM Flex System V7000 Storage Node. It provides an ability to mirror data between storage pools in the same IO group of the IBM Flex System V7000 Storage Node cluster. This capability can be used to protect critical data and increase availability, as well as to copy volumes for others to use for development, test and other application needs.

10.1 Volume mirroring and migration options

Volume mirroring is a function that is designed to allow a volume to remain online even when the primary storage pool backing it becomes inaccessible. This function is designed to protect the volume from storage infrastructure failures by providing seamless mirroring between different storage pools.

This function is provided by a specific volume mirroring function built into the I/O stack, and cannot be manipulated like the FlashCopy or other copy service functions. This feature does, however, enable a migration functionality, which can be obtained by splitting the mirrored copy from it's source or by using the **`migrate to`** command. Volume mirroring does not make use of or control external back-end storage mirror or replication functions as they are independent of the IBM Flex System V7000 Storage Node.

10.1.1 Tunable timeout

Generally, with this feature, the host IO completes when both copies of the volume mirror are written. This feature has a tunable latency tolerance. This is designed to provide an option to give preference to losing the redundancy between the two copies. This tunable timeout has two value settings, either *latency* or *redundancy* to chose from.

The *latency* tuning option (set by running **`svctask chvdisk -mirrorwritepriority latency`**) is the default and prioritizes host I/O latency. This setting produces a preference to host I/O over availability.

However, you might need to give preference to *redundancy* when availability is more important than I/O response time. To accomplish this task, run command **`svctask chvdisk -mirrorwritepriority redundancy`**.

Regardless of which option you choose, volume mirroring can provide extra protection for your environment and ease of management for moving data around to better serve your needs.

10.1.2 Usage of mirroring for migration

With regard to using mirroring for migration, there are several use cases that can exist:

- Creating multiple volume images using Volume Mirroring then Split into New Volume: This is one of the most common uses of this function. This option allows you to use the RAID 1 functionality to create two copies of data that initially have a set relationship (one primary and one secondary) but then break the relationship (both primary and no relationship) and make them independent copies of data. You can use this mode to migrate data between storage pools and devices. You might use this option if you want to move volumes to multiple different storage pools. You can only mirror one volume at a time.
- Volume migration using Move to Another Pool: This option allows any volume to be moved between storage pools without interruption to host access. This option is effectively a quicker version of clicking **Volume Mirroring** → **Split into New Volume**. You might use this option if you want to move volumes in a single step or do not already have a volume mirror copy.
- Migrating from external storage using Image mode: This option allows you to migrate data from an already existing volume or LUN on another external storage subsystem into the IBM Flex System V7000 Storage Node as an image mode mdisk. With this function, the volume can then be mirrored as a second copy to another already existing generic, striped

volume storage pool creating a copy without putting any metadata on the original device (Image mode mdisk); therefore, leaving the source data intact. After they are migrated in they are fully accessible as volumes for use by host and any of the other features of the IBM Flex System V7000 Storage Node.

Migration methods: When planning for use of the migration methods listed above, even though it may not be to distribute access, you will need to take a brief outage to install the host drivers for the IBM SAN Volume Controller. See *IBM System Storage SAN Volume Controller Host Attachment User's Guide*, SC26-7905 for detailed instructions. Make sure to consult the revision of the document that applies to the proper version of IBM SAN Volume Controller.

10.1.3 Managing Volume Mirror and migration with the GUI

To begin the Volume Mirror process you must first select the Volume by Pool from the storage Pool icon drop down list (Figure 10-1).

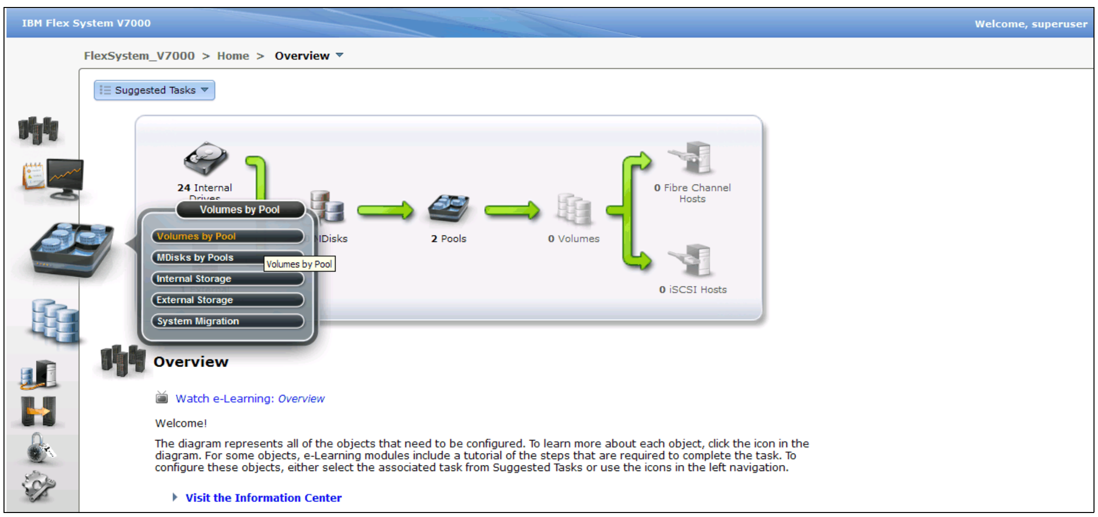


Figure 10-1 Volumes by Pool from the Storage Pool icon

To make a volume mirror using the GUI, you need to click **Add Mirrored Copy** from the **Volume Copy Actions** menu drop down box (Figure 10-2 on page 426).

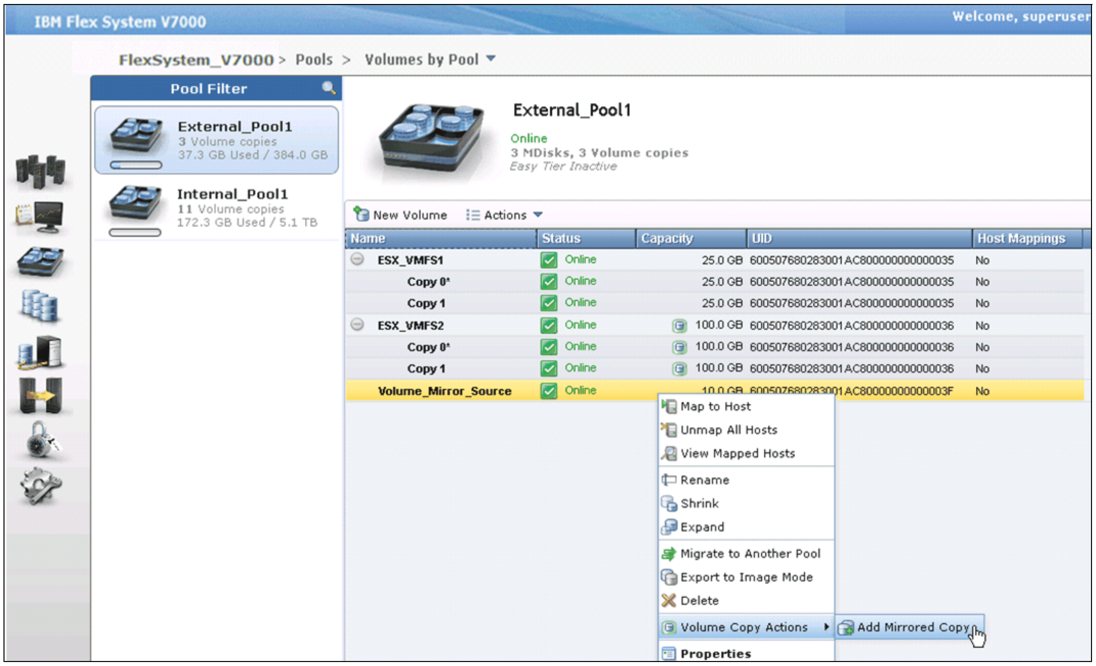


Figure 10-2 Add Volume Mirrored Copy

After you complete this action, you can specify the type of volume mirror, that is, generic or thin provisioned, and select the storage pool to use for the copy (Figure 10-3). Make sure that you select a storage pool with sufficient space and similar performance characteristics, and then click **Add Copy**.

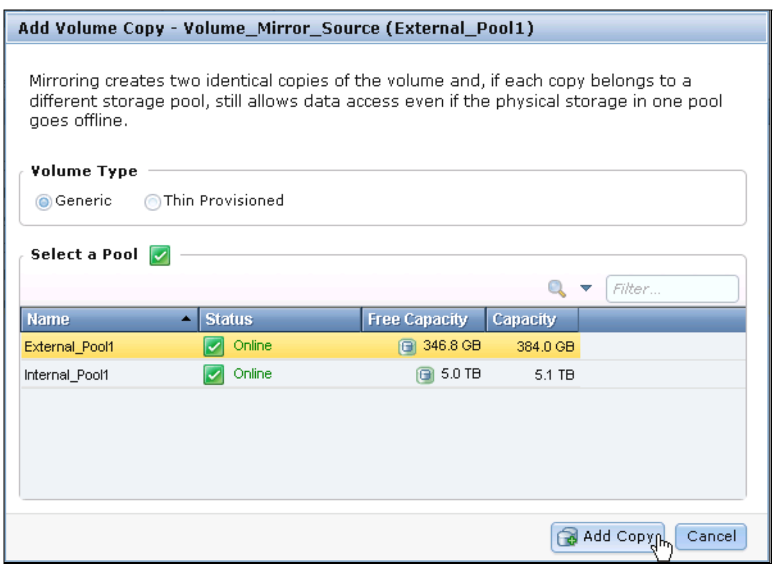


Figure 10-3 Confirm Volume Mirror type and storage pool to use for the mirror

After you create your mirror, you can view the distribution of extents on the assigned mdisks for each copy or for the entire volume (Figure 10-4). You can also view the mirroring progress percentage by high lighting the volume and selecting the *Volume Synchronization* tab in the Running Tasks radial (Figure 10-5).

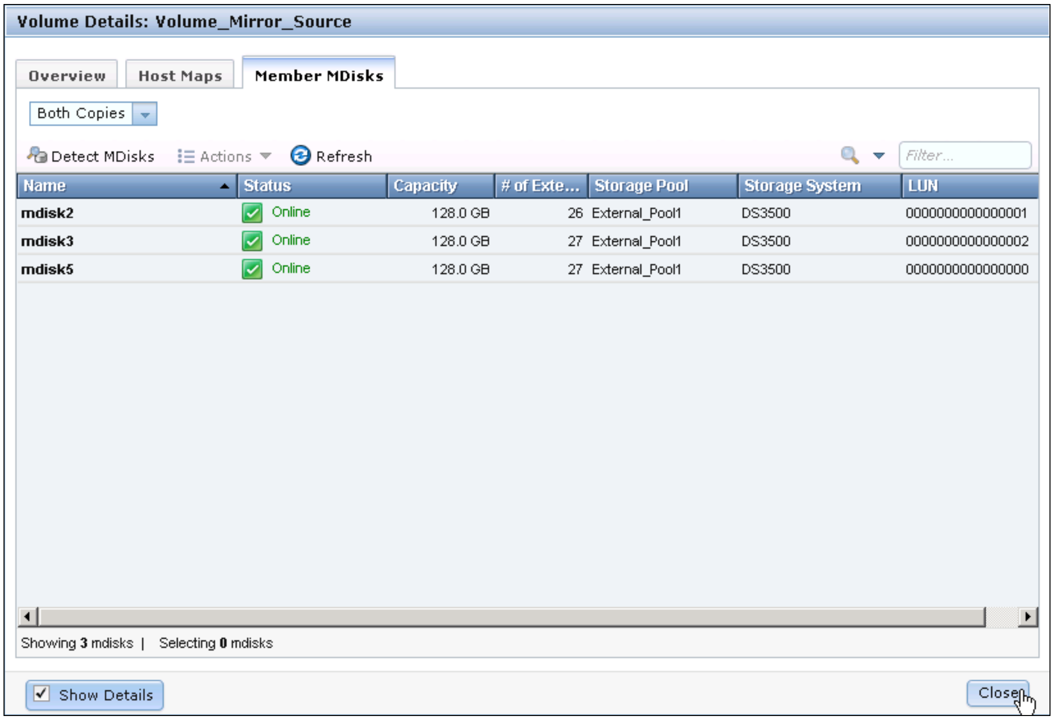


Figure 10-4 The distribution of extents for primary and mirror copy of a volume

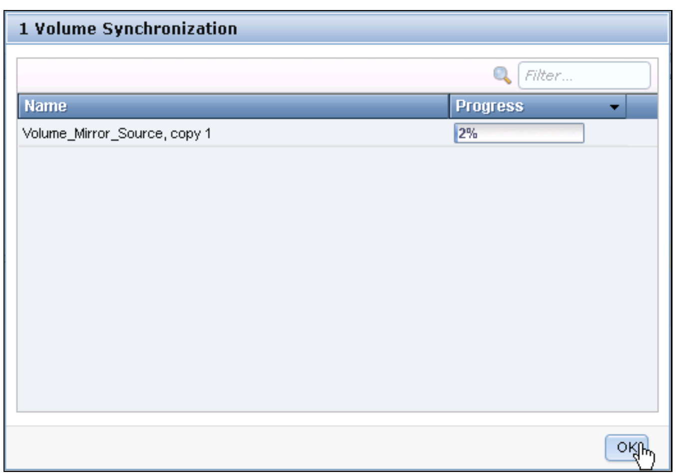


Figure 10-5 Progress of a mirror copy creation as viewed by using Running Tasks

Extent distribution: Extent distribution for the mirror copy is automatically balanced as well as possible within the storage pool selected.

After the copy completes, you have the option of splitting either copy of the mirror into a new stand-alone volume (Figure 10-6 on page 428).

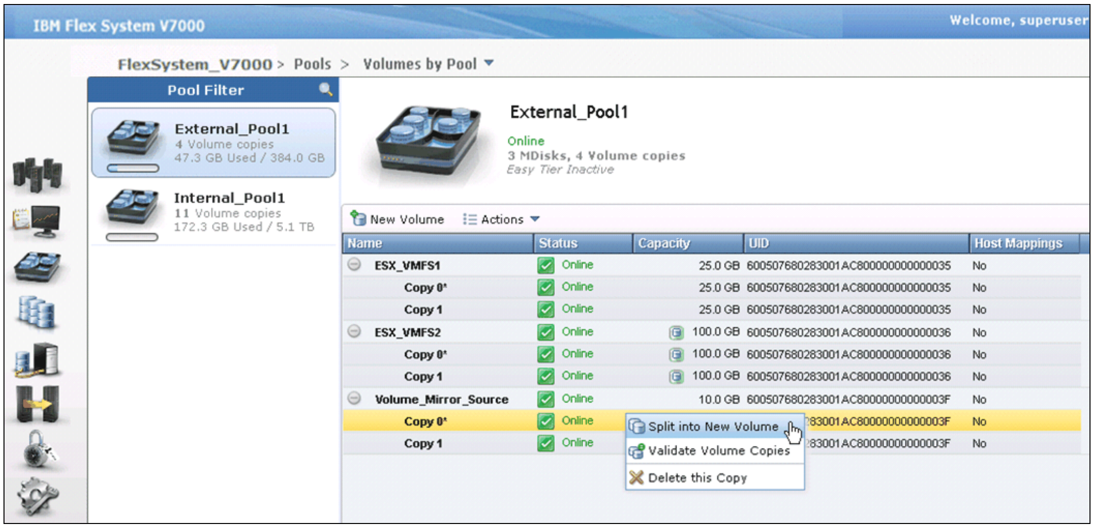


Figure 10-6 Selection of Split into New Volume

After you click **Split into New Volume** on either Copy0 or Copy1, you can specify a new volume name and confirm the split (Figure 10-7).

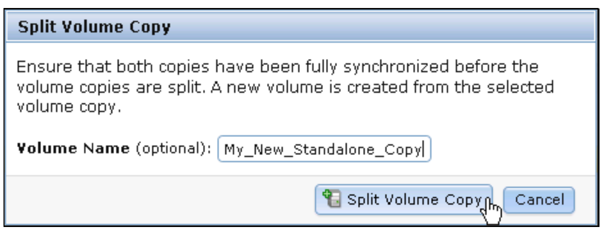


Figure 10-7 Confirmation of volume mirror split

After providing a new volume name (optional) and confirming the split, you can see the results as shown in Figure 10-8.

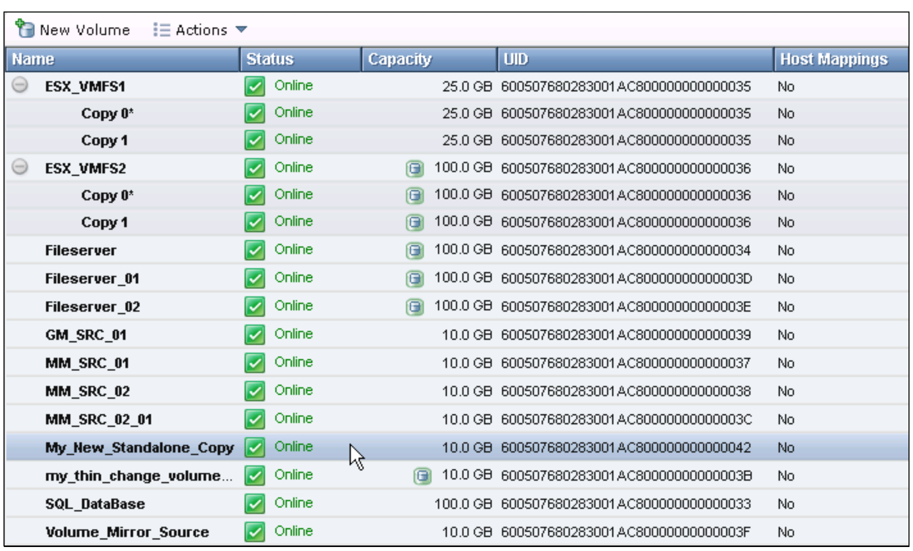


Figure 10-8 Results of volume mirror split

Splitting volume copies: When you split a volume copy, the view returns to the pool in which it was created, not to where the primary copy exist.

If you want to migrate your volumes to another storage pool in one step instead of two, you can use the **Migrate to Another Pool** option (Figure 10-9).

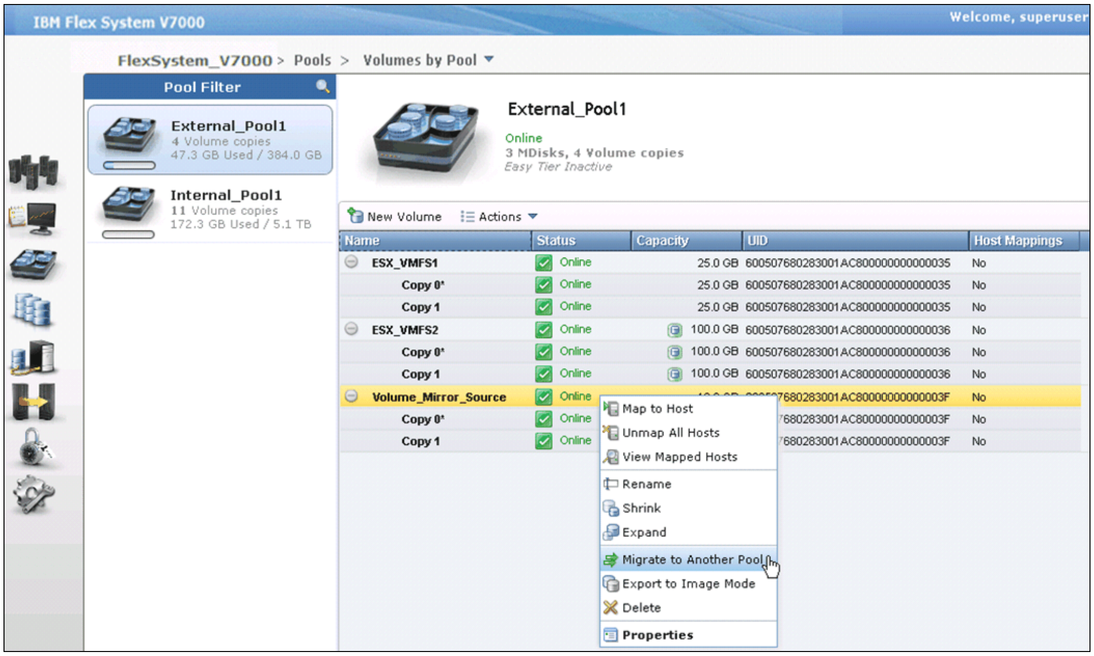


Figure 10-9 Using the Migrate to Another Pool option

Migration limitation: You cannot migrate more than one volume at a time. For this reason, Copy Services functions can be more expedient, if available.

If the volume has only one copy, a storage pool selection window opens. If it has two, the same window with a variation that allows you to choose the copy to migrate opens as shown in Figure 10-10.

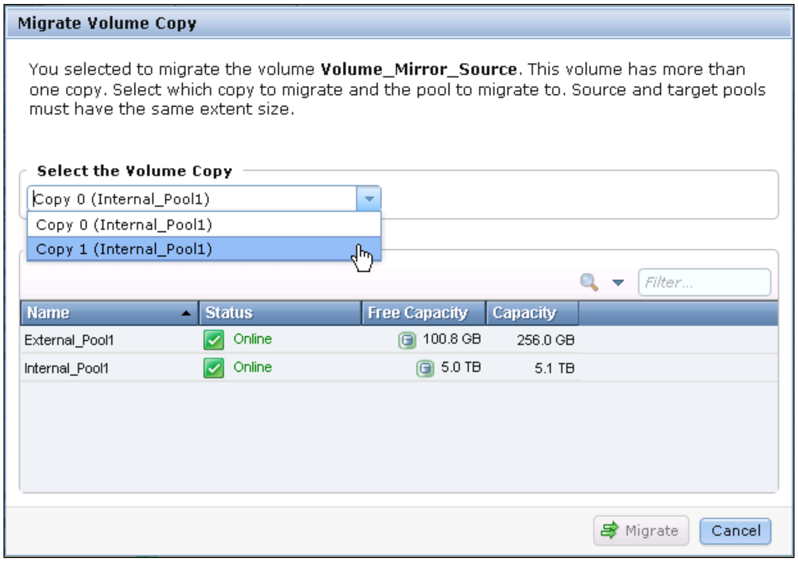


Figure 10-10 Selecting the destination storage pool of a mirrored volume

The selection you are presented within Figure 10-10 denotes the current pool of each volume copy, so you can better determine which storage pool to use.

Finally, you can also migrate a volume that is from an external storage system and you desire to migrate the volume's data to an internal storage pool. For details on steps to follow to perform this operation, refer to Chapter 7, "Migration wizard" on page 265.

11

SAN Connections and Configuration

This chapter discusses the types of connectivity of IBM Flex System V7000 Storage Node to the compute nodes in the IBM Flex System Enterprise Chassis for storage provisioning. It includes the three types of storage area network - fibre channel (FC), iSCSI, and FC over Ethernet (FCoE). To do this we will look at the configuration of two of the I/O Module models IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch and IBM Flex System FC3171 8Gb SAN Switch. Also, we look at how the IBM Flex System V7000 Storage Node can connect to external storage controllers and SANs in order to manage.

11.1 Storage Area Network Overview

Storage Area Network (SAN) has long been regarded as consisting of dedicated equipment which connects host servers to share remote storage. This has enabled storage of varying size and performance to be used and allocated more efficiently and help eliminating islands of storage. Multiple servers of differing platform types can benefit from a single new storage controller and its associated attributes such as performance or provisioning technologies. With the increased efficiencies and economies this has brought, organizations have been well prepared for the storage capacity increase and demand experienced in recent years. The most common and reliable protocol used in today's SAN has been Fibre Channel (FC).

Fibre Channel SANs are generally regarded as the high-performance approach to storage networking. With a Fibre Channel SAN, storage arrays are equipped with FC ports that connect to FC switches. Similarly, servers are equipped with Fibre Channel host bus adapters (HBAs) that also connect to Fibre Channel switches. Therefore, the Fibre Channel SAN fabric which is the set of FC switches, are a separate and exclusive network for storage traffic. FC offers relatively high-speed, low-latency, and back-pressure mechanisms that provide lossless behavior. It is designed not to drop data packets during periods of network congestion.

The speeds offered from FC has until recently always been faster than that of Ethernet where FC used speeds of 2 Gbps, 4 Gbps, 8 Gbps and more recently16 Gbps, Ethernet has offered 100 Mbps or 1 Gbps. But with improved and faster Ethernet communications to a host server 10 Gbps is becoming more widely available and used for host server connections. Higher speeds of 40-Gbps Ethernet are becoming available, and a 100-Gbps Ethernet will soon arrive. With newer 'lossless' Ethernet networks also, from a simple data transmission speed perspective, Ethernet can now meet or exceed the speeds that are available by using FC.

With these increased Ethernet transmission speeds becoming available there is a growing interest in convergence where FC and traditional Ethernet data traffic can be merged and transferred over a single medium. Another driving factor for convergence is improved asset utilization and cost of ownership, similar to the convergence of voice and data networks that has occurred. By using a single infrastructure for multiple types of network traffic, the costs of procuring, installing, managing, and operating the data center infrastructure can be lowered. Where multiple types of adapters, switches, and cables were once utilized for separate networks, a single set of infrastructure will take its place, providing savings in equipment, cabling, and power requirements. The improved speeds and capabilities of a lossless 10-Gbps Ethernet offer the hope of a converged environment.

This section looks at how the IBM Flex System V7000 Storage Node is able to utilise convergence technologies in order to communicate with the compute nodes and external storage equipment. We look at the three type of SANs offered within the IBM Flex System Enterprise Chassis iSCSI, FCoE and FC.

For in depth details of FCoE and iSCSI including protocols which can use the high speed Ethernet networks and recommendations see *Storage and Network Convergence Using FCoE and iSCSI*, SG24-7986. For in depth details of FC see *Introduction to Storage Area Networks and System Networking*, SG24-5470.

11.2 Connection to Chassis I/O Modules

The IBM Flex System V7000 Storage Node uses 10Gb Ethernet adapter for iSCSI and FCoE communications plus optional fibre channel (FC) adapters for FC connectivity to the compute

nodes. As stated in 2.4.1, “Control canister” on page 48, each canister on the IBM Flex System V7000 Storage Node has host attachments on one control canister which match the configuration of the second canister and must use the same slots. The Network Adapter slots correspond to the following I/O module slots on the IBM Flex System Enterprise Chassis:

- Network Adapter slot 1 in each node canister connects to IO modules 1 and 2
- Network Adapter slot 2 in each node canister connects to IO modules 3 and 4

Network Adapter cards must be installed in pairs (one card per node canister) and the following configurations are supported:

- Two or four 10Gb CNA cards
- Two 8Gb FC cards
- Two 10Gb CNA cards and two 8Gb FC cards

Adapters Supported on IBM Flex System V7000 Storage Node include:

- IBM Flex System EN2024 4-port 1Gb Ethernet Adapter
- IBM Flex System EN4054 4-port 10Gb Ethernet Adapter
- IBM Flex System FC3172 2-port 8Gb FC Adapter
- IBM Flex System FC5022 2-port 16Gb FC Adapter
- IBM Flex System FC3052 2-port 8Gb FC Adapter (Emulex)
- IBM Flex System CN4054 10Gb Virtual Fabric Adapter
- IBM Flex System CN4058 8-port 10Gb Converged Adapter

The Network Adapter on IBM Flex System V7000 Control Enclosure canister will communicate to switch modules on the IBM Flex System Enterprise Chassis I/O module ports. Some pass-through modules are available as an alternative to switches on the IBM Flex System Enterprise Chassis I/O modules. I/O Modules supported on IBM Flex System V7000 Storage Node include:

- IBM Flex System EN2092 1Gb Ethernet Scalable Switch
- IBM Flex System EN4091 10Gb Ethernet Pass-thru
- IBM Flex System Fabric EN4093 10Gb Scalable Switch
- IBM Flex System Fabric EN4093R 10Gb Scalable Switch
- IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch
- IBM Flex System FC3171 8Gb SAN Switch
- IBM Flex System FC3171 8Gb SAN Pass-thru
- IBM Flex System FC5022 16Gb SAN Scalable Switch

For more information of these modules, refer to 1.5, “I/O Modules” on page 23.

The I/O module will determine the connectivity type of the IBM Flex System V7000 Storage Node to the compute nodes. The storage node canister must have an appropriate Network Adapter suitable to the I/O module in order to communicate with the required protocol.

Each of the two node canisters in the IBM Flex System V7000 Storage Node can have up to two Network Adapter cards. These will connect to the I/O modules so that the Storage Node can communicate with and present storage to the compute nodes.

Figure 11-1 on page 434 shows the logical connections from the Storage Node canister Network Adapters to the I/O module switches. Each logical connection shown in the figure represents four serial links to an I/O module switch. The Network Adapter port uses only one of these four ‘lanes’ per switch, the unused lanes are provided for future options. Each of these individual I/O links or lanes can be wired for 1Gb or 10Gb Ethernet, 8 or 16 Gbps Fibre Channel.

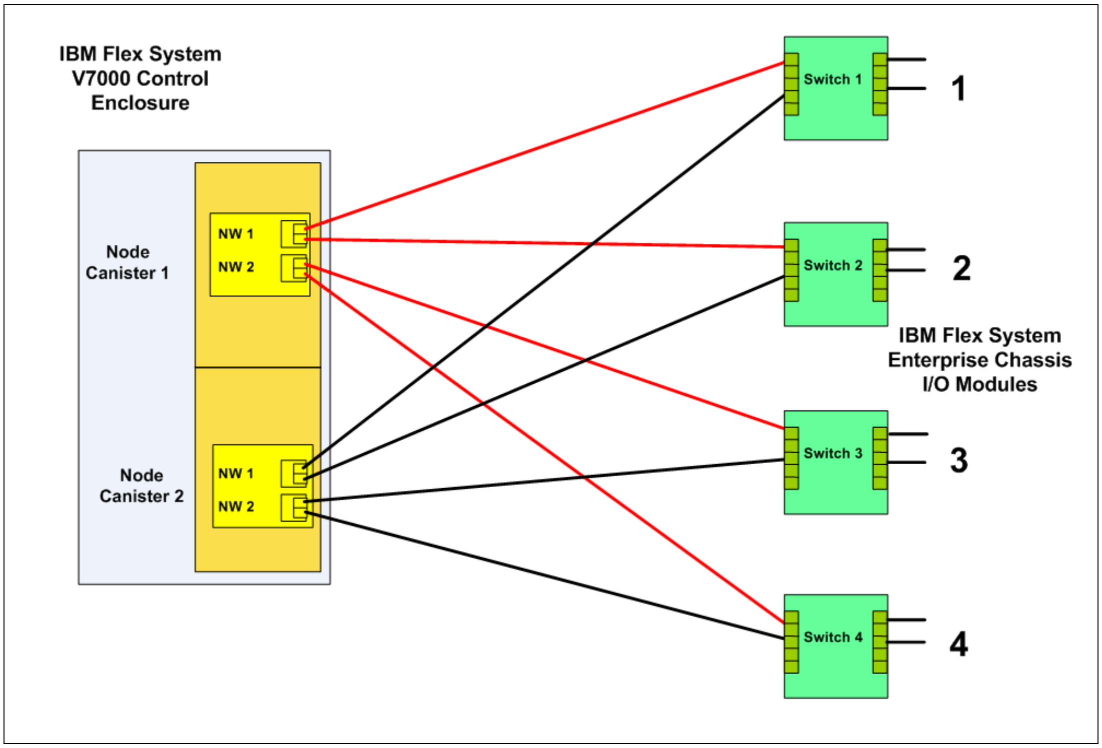


Figure 11-1 Logical layout of Storage Node to I/O module switch interconnects

I/O Module connection summary

The storage node canister can communicate with all 4 of the I/O switch modules in the IBM Flex System Enterprise Chassis via both of the canister Network Adapters. NW 1 on each canister will be used for Ethernet communications via I/O module switches 1+2 where traffic can be iSCSI on a TCP/IP network, or FCoE thus providing a network convergence capability. NW 2 can be used for either an extra Ethernet attachment to I/O module switches 3+4 or can be used for FC connection to compute nodes and external FC fabrics for external storage attachment.

Figure 11-2 on page 435 shows the SAN area which will be examined in this section, it will detail connections within the IBM Flex System Enterprise Chassis predominantly but will also look at connectivity to external storage and configuration of the I/O switch modules.

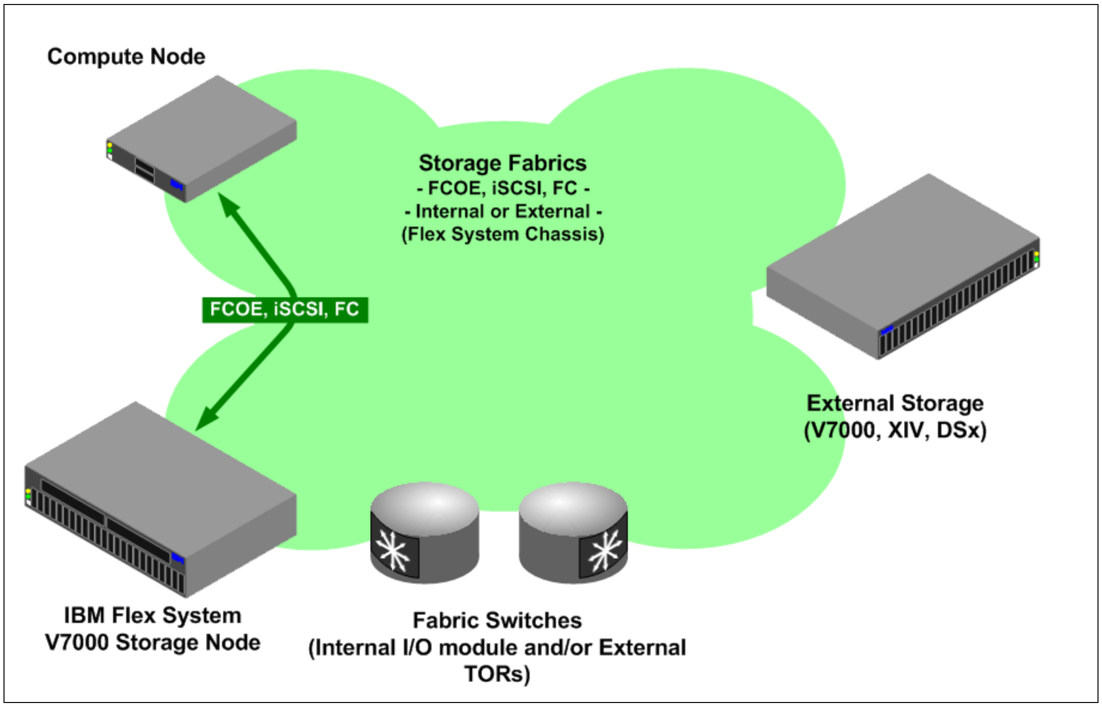


Figure 11-2 SAN Connectivity within the IBM Flex System Enterprise Chassis

11.3 iSCSI Connectivity

The Internet Small Computer System Interface (iSCSI) protocol is the encapsulation of the industry standard SCSI protocol within TCP/IP packets. The iSCSI protocol provides a block-level storage capability similar to Fibre Channel (FC) storage area network (SAN) technology. This is essentially the same system of encapsulating the SCSI protocol within an external “carrier.” The difference is that the iSCSI SAN uses Ethernet instead of FC transport technology. The iSCSI protocol uses TCP to transport SCSI I/O commands over a network. It provides block-level data access where it encapsulates SCSI commands into protocol data units (PDUs) within the TCP/IP protocol, and then transports them over the network to the target device.

iSCSI transactions occur between an iSCSI initiator (hardware or software) that transmits a request (such as read/write) and an iSCSI target. This iSCSI target processes the request and responds with the appropriate information, such as data and sense. The initiating device in the Flex System environment is the compute node and the target devices are the IBM Flex System V7000 Storage Node canisters. The connectivity between the initiator and target is shown in Figure 11-1 on page 434 and detailed in that section. The type of connectivity would be determined also by the compute node platform, operating system and its device drivers.

11.3.1 Session establishment and management

A *session* is the creation of a communication tunnel from the iSCSI initiator to the iSCSI target. An iSCSI session consists of an *iSCSI login phase* and a *full feature phase*. At least one session between the initiator and the target must be enabled through an iSCSI login process. A login protocol data unit (PDU) is used to negotiate any variable parameters between the two parties and can start a security routine to authenticate permissible connectivity. When the login is successful, the target issues a login success to the initiator.

Otherwise, it issues an unsuccessful login. iSCSI can support multiple connections per session (MCS) to increase aggregate bandwidth or, for several links, to improve reliability.

PDU structure

A *protocol data unit* is the basic message packet that travels between a client and target. It consists of a basic header segment (BHS) and additional headers segments (AHS). It also includes optional parameters, such as cyclic redundancy check (CRC) data segments and data digests.

iSCSI error handling

An IP network is susceptible to the high probability of errors in data delivery. The iSCSI protocol provides several measures to counter potential errors. The design requires iSCSI to perform its functions within a Internet Protocol network safely and use appropriate quality of service (QoS) procedures.

iSCSI security

The iSCSI can be used in networks where unauthorized data can be accessed, allowing for different security methods. Encoding means, such as IPSec, which use lower levels, do not require additional matching because they are transparent for higher levels, and for iSCSI. Various solutions can be used for authentication, for example, CHAP, Kerberos, or private keys exchange. An iSNS server can be used as a repository of keys.

11.3.2 iSCSI initiators

iSCSI initiators can be implemented by software or hardware. Software initiators can be augmented by TCP-offload Ethernet adapters.

Software-only solutions

Software initiators and targets are virtual SCSI adapters that are written as part of the operating environment. They use the processor resources and network adapters of the host to transfer data. Software endpoints are easy to deploy and are low-cost or free with the host operating system.

Software implementations can drive higher throughput than other implementations if sufficient host processor resources are available. This higher throughput is especially true of cases where smaller block sizes are used. Integration with the host operating system usually works well, by using existing management tools and interfaces. Starting a host from an iSCSI device is not possible when using software initiators unless a pre-startup execution environment exists. At a minimum, a DHCP server and a file transfer protocol, such as Trivial File Transfer Protocol (TFTP), are required.

Figure 11-3 compares the iSCSI initiator technologies.

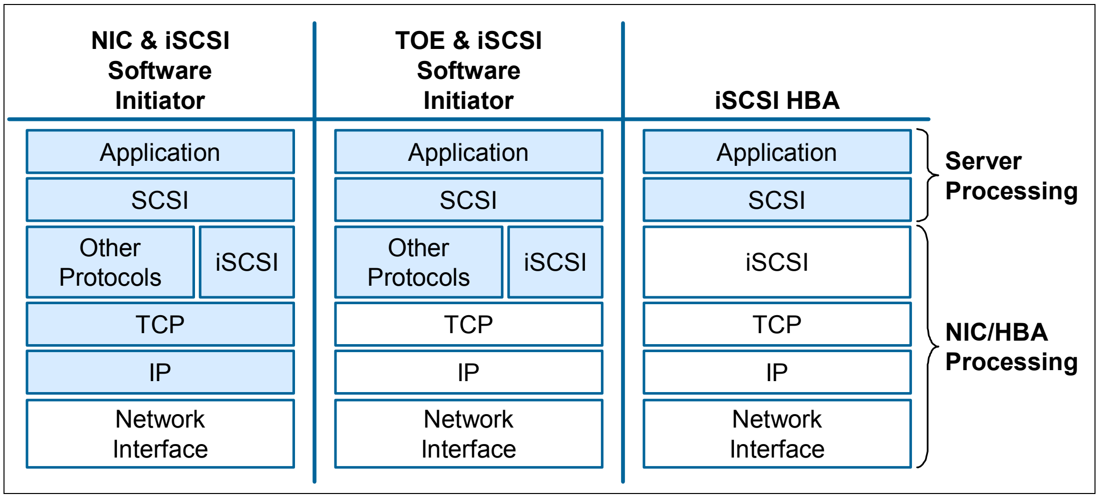


Figure 11-3 Comparison of iSCSI initiator options

Software with hardware assistance

Hardware assistance in the context of an iSCSI software endpoint generally comes in the form of a TCP offload engine (TOE). With a TOE, the TCP stack processing, including framing, reordering of packets, check sums, and similar functions, are offloaded to a dedicated card with its own network interface port. The TOE card can be a general-purpose card that can offload TCP traffic, or it can be restricted to just accelerating iSCSI traffic.

TOE adapters benefit most from software initiators with additional host processor offload. TOE adapters can also support advanced networking features such as link aggregation. Because the software initiator is still used on the host, integration with layered management applications is unaffected by the addition of the TOE hardware.

Hardware-only solutions

Hardware iSCSI adapters (hardware initiators) offload TCP stack processing functions and iSCSI command processing functions. The hardware adapter looks and functions similar to a SCSI disk interface, just as an FC HBA does. The operating system has no knowledge of the underlying networking technology or interfaces. A separate management interface is used to configure the networking parameters of the card.

Hardware-only solutions offload the largest amount of processing from the host processor. Because they function as SCSI adapters, you can start an operating system from them if they provide the appropriate host BIOS interfaces and are recognized as a startup device. Advanced networking features might not be available because of software visibility to the network functions in the card.

The IBM Flex System V7000 Storage Node and some compute node adapters use iSCSI hardware-only initiators. Other compute node adapters do have TOE capability.

<< Author Comment - Need to verify last statement as some doc did mention TOE as part of specs >>

Figure 11-4 on page 438 compares the features of iSCSI adapter options.

	NIC	TOE	iSCSI HBA
Definition	Provides Ethernet connectivity	A specialized NIC that provides additional functionality	An HBA that provides Ethernet connectivity and additional functionality
Offloaded burden	Physical and data link communication	TCP/IP, physical and data link communication	iSCSI read/write processing, TCP/IP processing, physical and data link communication
Burden managed by server CPU	iSCSI protocol management and TCP/IP protocol management	iSCSI protocol management	None
Require software-based initiator	Yes	Yes	No
iSCSI performance	Adequate	Good	Best
Cost	$	$$	$$$$

Figure 11-4 Comparison of iSCSI adapter technologies

11.3.3 iSCSI Multisession Configuration and Support

Multisession support, or host-level multipath support, provides multiple paths between iSCSI initiators and targets, which is used for high availability and load balancing.

Some products require that an iSCSI initiator log in with a specific IP address of the target versus logging in with a target name. Logging in with a target name in those products causes the iSCSI initiator to log into all of the IP addresses in the target and causes session reinstantiation that results in session loss of the previous log in. Flex System V7000 Storage Node offers single subnet multisession configurations and multiple and dual subnet multisession configurations that remove the restriction against logging into a target name. iSCSI initiators use Internet storage name service (iSNS) servers to log in, which enables the discovery of the iSCSI targets and their IP addresses.

An iSCSI session is a TCP relationship between an iSCSI initiator node port and an iSCSI target node port. Once established, iSCSI control, data, and status messages are communicated over the session.

Each session is identified with multiple iSCSI initiator names and the target name plus the iSCSI initiator session ID (ISID) plus the target portal group tag (target-side ID).

The Flex System V7000 Storage Node multisession approach is to have relationships between multiple initiator instances (separate ISIDs) to a single target portal group in an iSCSI target. Such an approach is similar to having multiple SCSI initiator ports from the same iSCSI initiator device have relationships with a single SCSI target port.

The following points apply to multisession configuration:

- The AIX iSCSI initiator does not support multisession.
- The use of the supported iSCSI multipathing is recommended when running the IO with multisession.
- If you use CHAP authentication under the Microsoft Windows operating system, use the same secret across each session that is initiated from that host to a particular target.

- If the Microsoft Windows clustered system is rebuilt or restored from a backup, hosts must log off and then log back on to any previously connected iSCSI targets that are supplied from the system.
- If the host application uses SCSI 3 Persistent Reservation, you must run only a single iSCSI session.
- Running with multiple iSCSI sessions is not supported for applications using Microsoft 2003 MSCS and Microsoft 2008 Failover clustering over an iSCSI interface; you must run only a single session in these situations.

11.3.4 iSCSI Multipath Connectivity

You can choose from several technologies to improve the availability of iSCSI by using multiple network paths. Multipathing requires a software layer above the iSCSI layer to make sense of the multiple physical paths between the initiator and target and to manage them appropriately.

Within the IBM Flex System Enterprise Chassis I/O modules 1+2 will be populated by Ethernet switches, these will allow iSCSI connectivity between a compute node and the IBM Flex System V7000 Storage Node. It is important to configure for switch redundancy by using iSCSI multipathing device driver software on the compute node host operating system. At the IBM Flex System V7000 Storage Node canisters both iSCSI Ethernet ports should be used.

In Figure 11-1 on page 434 it can be seen that NW 1 on each node canister is connected to Switch 1 and Switch 2. When configuring for iSCSI connectivity as shown in Figure 11-5, port 2 is represented by the link to Switch 1 and port 3 is represented by the link to Switch 2 across both canisters.

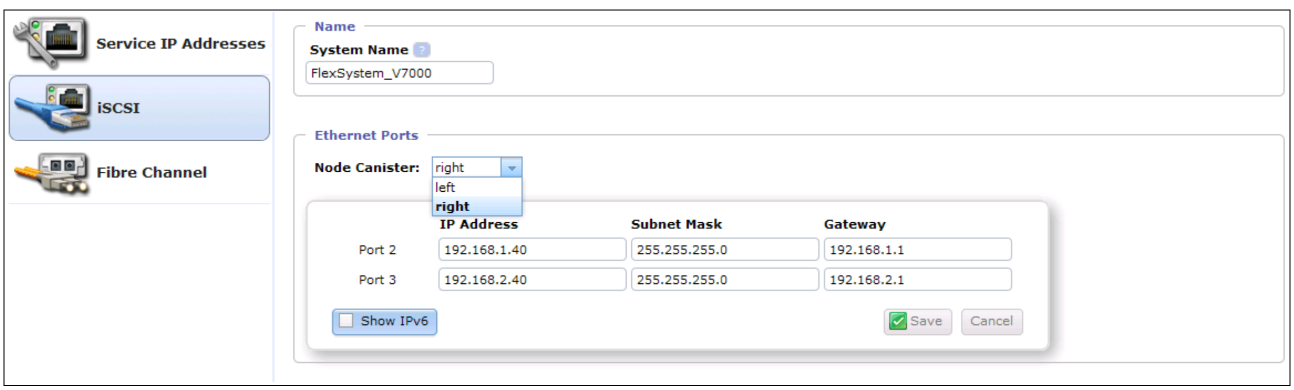


Figure 11-5 iSCSI ports 2 and 3 linking to switches in IBM Flex System Enterprise Chassis

At the host 4 paths can be configured to the IBM Flex System V7000 Storage Node using each of the 2 Ethernet ports from each of the storage node canisters. In Figure 11-6 on page 440 the example shows that both switches are used with separate networks or subnets, each communicating with a storage node canister and a compute node.

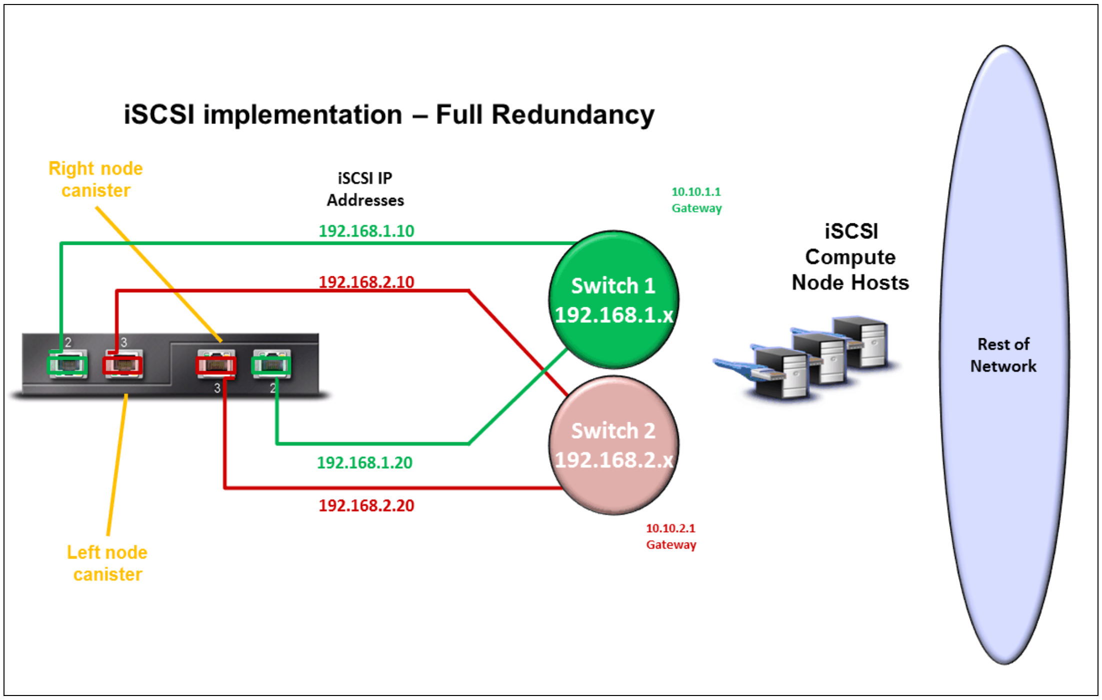


Figure 11-6 iSCSI redundancy using both Ethernet switches

11.3.5 Configuring Multiple iSCSI Host Links

This section will show how to configure the IBM Flex System V7000 Storage Node canisters for multiple connections to a x240 compute node running MS Windows. The example will create the configuration as shown in Figure 11-6, the x240 compute node has two Ethernet adapter ports which are already configured with IP addresses. This example does not take into account of network planning required or configuration on the Ethernet switch or TCP/IP network being used. It is recommended that a form of access limitation is implemented across the TCP/IP network such as use of VLAN and network security, the example below does not include these details.

Refer to 6.2.2, “Create an iSCSI attached host” on page 259 for details on how to define an iSCSI host on theIBM Flex System V7000 Storage Node.

Figure 11-7 on page 441 shows the iSCSI network view of the IBM Flex System V7000 Storage Node configuration GUI. From here the IQN can be seen and where IP address for each canister can be set:

1. Navigation pane to present view Node_ID Settings Network
2. Canister select option
3. IP address of Port 2 and Port 3 to be changed
4. Node and IQN used by the host when initiating the connection

Table 11-1 on page 441 shows the IP configuration settings for each iSCSI port on both canisters, as described earlier are attached to Ethernet switches in IBM Flex System Enterprise Chassis I/O Module ports 1 and 2.

Table 11-1 IP port configuration on *IBM Flex System V7000 Storage Node canisters*

Canister	Ethernet Port	IP Address
left	Port 2	192.168.1.10
left	Port 3	192.168.2.10
right	Port 2	192.168.1.20
right	Port 3	192.168.2.20

The host is detailed in section 6.2.2, “Create an iSCSI attached host” on page 259 where the IQN of the x240 compute node is defined. The IQN number is taken from the host as provided by the operating system.

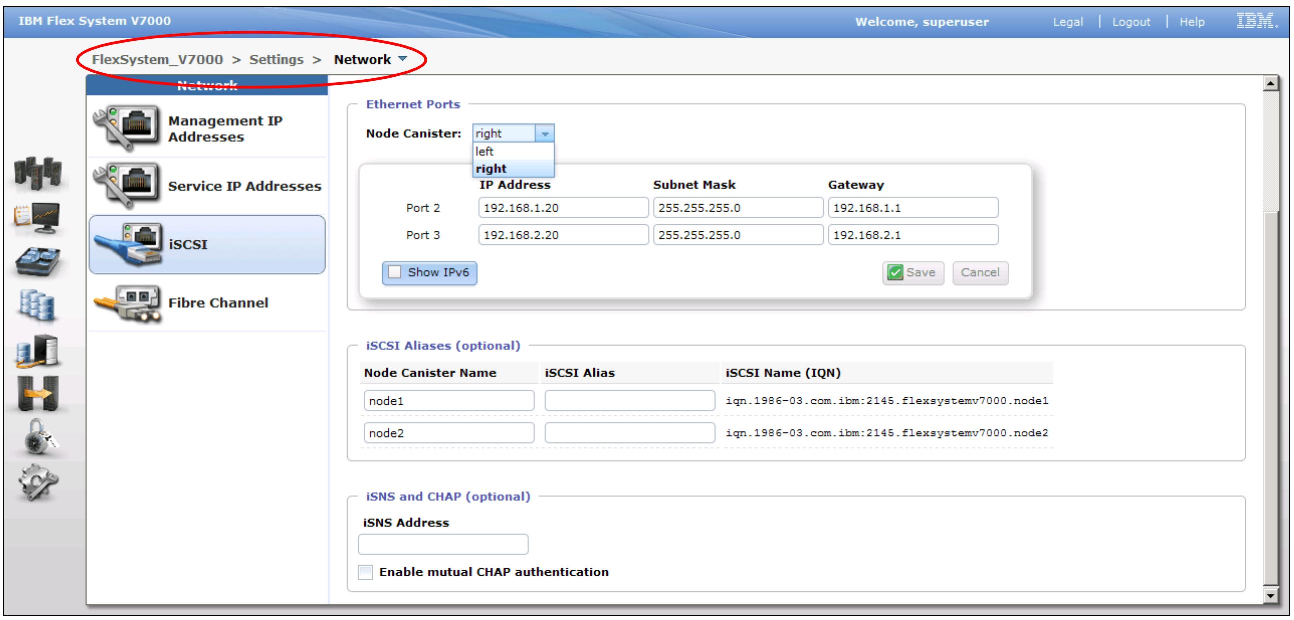


Figure 11-7 Setting up the iSCSI ports on the node canisters

The IQN for each IBM Flex System V7000 Storage Node canister is:

- iqn.1986-03.com.ibm:2145.flexsystemv7000.node1
- iqn.1986-03.com.ibm:2145.flexsystemv7000.node2

These are used by the host to connect to the IBM Flex System V7000 Storage Node.

When the canister ports are configured as shown in Table 11-1 the link should be tested by the host. We used a ‘ping’ to check connectivity across the switch.

The switching involved for iSCSI is at the TCP/IP network level where multiple paths are planned to eliminate a single point of failure in the network and for load balancing. This is also determined by the iSCSI device driver of the host. Planning for connectivity between iSCSI source and target devices would be at the TCP/IP network. The node canister ports are assigned IP addresses and TCP/IP connectivity to them is tested from each host.

The full configuration process for iSCSI attachment to a WIN8 host is detailed in 12.3.2, “Windows 2008 R2 iSCSI volume attachment” on page 480.

11.4 FCoE Connectivity

Fibre Channel over Ethernet (FCoE) is a method of sending Fibre Channel frames directly over an Ethernet network. It relies on a new Ethernet transport with extensions that provide lossless transmission of storage data. The new Ethernet transport standard is the Fibre Channel - Backbone - 5 (FC-BB-5) standard which specifies that FCoE is intended to operate over an Ethernet network that does not discard frames in the presence of congestion. Such an Ethernet network is called a *lossless Ethernet* in this standard.[1] The standard also states that devices shall guarantee in-order delivery of FCoE frames within the Lossless Ethernet network.[2]

This is achieved by using a new set of Data Center Bridging (DCB) standards that enhance existing 802.1 bridge definitions. The enhancements provide a converged network that allows multiple applications to run over a single physical infrastructure. These DCB standards are intended to apply to more broader type of traffic than just for FCoE.

The DCB standards include:

- Priority-based Flow Control–IEEE 802.1Qbb (PFC)
- Enhanced Transmission Selection–IEEE 802.1Qaz (ETS)
- Congestion Notification–IEEE 802.1Qau (CN)
- Data Center Bridging Capabilities Exchange–IEEE 802.1Qaz

There have been a number of terms to describe these DCB standards but the term *Converged Enhanced Ethernet* (CEE) is now generally accepted by IBM and several other vendors in the T11 working group. The official IEEE 802 term is now *Data Center Bridging*.

FCoE Protocol stack

The FCoE requirement is the use of a lossless Ethernet, for example one that implements the DCB extensions to Ethernet as described above.

The structure of FCoE is that the upper layers of FC are mapped onto Ethernet, as shown in Figure 11-8 on page 443. The upper layer protocols and services of FC remain the same in an FCoE deployment. For example, zoning, fabric services, and similar functions still exist with FCoE. The difference is that the lower layers of FC are replaced by lossless Ethernet. Therefore, FC concepts, such as port types and lower-layer initialization protocols must also be replaced by new constructs in FCoE. Such mappings are defined by the FC-BB-5 standard.

[1] American National Standards Institute, Inc., *Fibre Channel — Fibre Channel Backbone - 5 (FC-BB-5)*, 4.4.4, “QoS and bandwidth” on page 26 at `http://fcoe.com/09-056v5.pdf`

[2] ibid (FC-BB-E page 27)

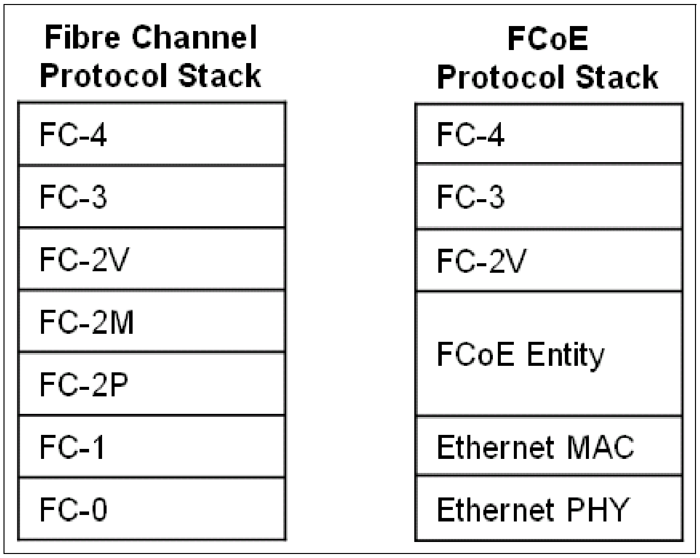


Figure 11-8 FCoE protocol mapping

Figure 11-9 on page 444 shows another perspective on FCoE layering compared to other storage networking technologies. The FC and FCoE layers are shown with the other storage networking protocols along with iSCSI.

In general, an FCoE network contains servers, lossless Ethernet (DCB-capable) switches, Fibre Channel Forwarders (FCFs) that provide FC fabric services, and storage devices. An existing FC SAN may or may not be present. With compute node connectivity to IBM Flex System V7000 Storage Node, the connection link is via I/O module lossless Ethernet FCF switches, a connected FC SAN does not have to be present.

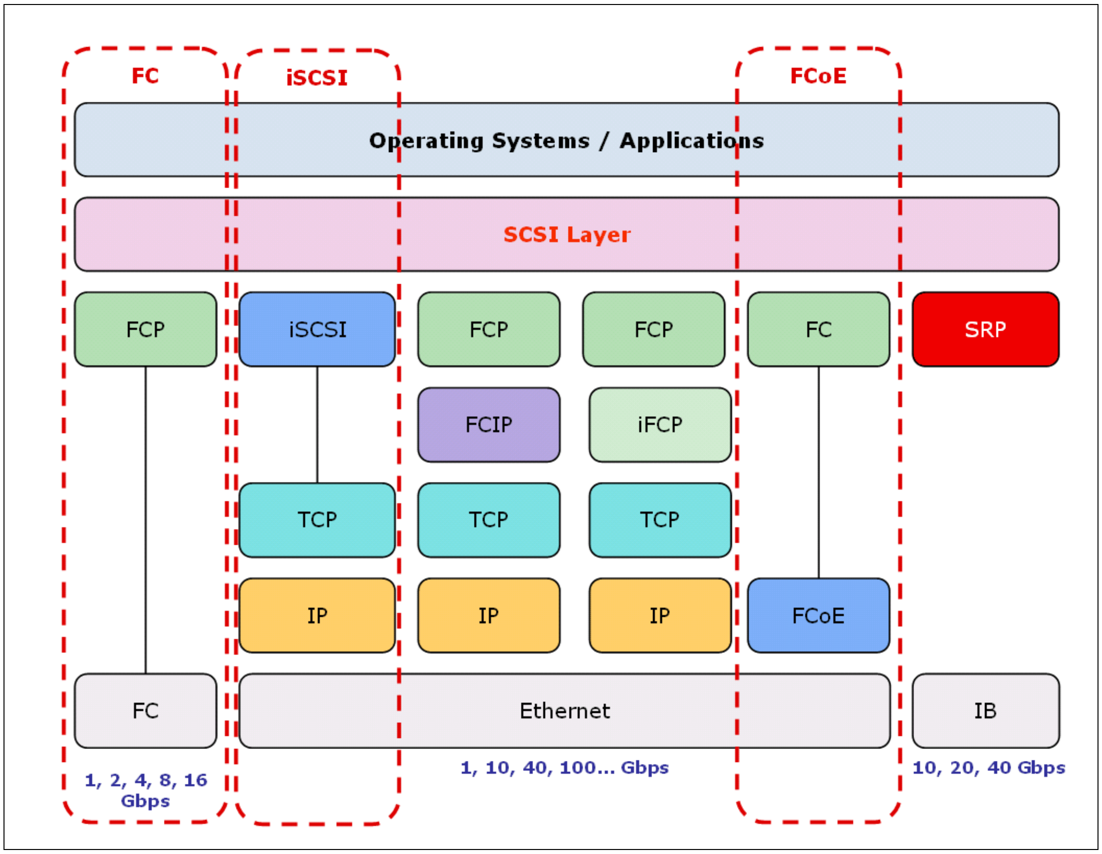


Figure 11-9 Storage Network Protocol Layering

11.4.1 Converged Network Adapters

In a traditional FC storage area network (SAN) and LAN environment, the server has an Ethernet adapter for data traffic and a Fibre Channel host bus adapter (HBA) for storage traffic. With converged networks which may include FCoE, these two adapters are replaced by a converged network adapter (CNA) that services both protocol stacks. A cable from the CNA connects to a lossless Ethernet switch, which eventually provides connectivity to an FCF and storage devices. Figure 11-10 on page 445 illustrates a CNA in an IBM Flex System Enterprise Chassis compute node server. In the compute node the only CNA adapter available is the IBM Flex System CN4054 10Gb Virtual Fabric Adapter.

Figure 11-10 on page 445 shows an example of FCoE connectivity of a compute node to the IBM Flex System V7000 Storage Node where the I/O module switch used is a IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch, this provides both FCF and DCB functionality.

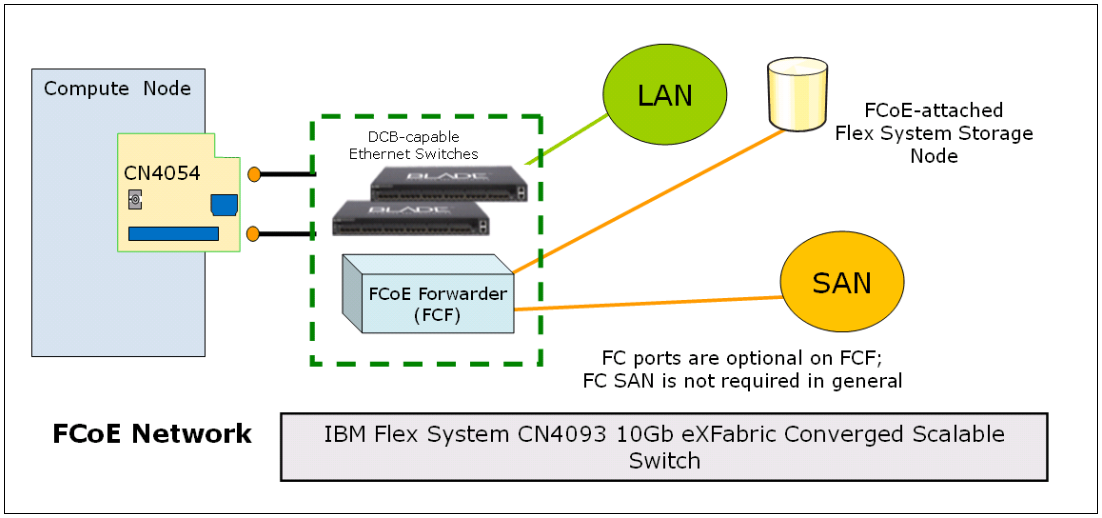


Figure 11-10 Compute node with CNA and FCF I/O module switch

Fibre Channel Forwarders

The IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch has an optional Fibre Channel Forwarder. The FCF function is the FC switching element in an FCoE SAN. It provides functions that are analogous to the functions provided by the FC switch in a traditional FC SAN. The most basic function is the forwarding of FCoE frames received on one port to another port based on the destination address in the encapsulated FC frame. The FCF is also the entity that handles FC logins, routing, zoning, and other FC services. As shown in Figure 11-8 on page 443 the lower layers of FC have changed with FCoE, but the upper layers are intact. The forwarding of FCoE frames between a compute node and an IBM Flex System V7000 Storage Node are therefore contained within the IBM Flex System Enterprise Chassis where the CN4093 provides the FCF switching functionality.

The CN4093 switch with its FCF function and FC ports can connect to external Fibre Channel SANs. In this case, the CN4093 10Gb Ethernet switch provides a gateway device function between FCoE and FC, transmitting frames between the two types of networks. As shown in Figure 11-10 the IBM Flex System V7000 Storage Node can manage external storage controllers using this capability to attach to FC SAN fabrics.

11.4.2 FCoE port types

In an FCoE network, virtual links are used across the lossless Ethernet network in place of the physical links in the FC network. The server negotiates a connection to the FCF device across the Ethernet network by using the FCoE Initialization Protocol (FIP). The server end of this connection is called a *VN_Port*. The FCF end is called the *VF_Port*. Two FCFs can also negotiate a connection across the Ethernet network, in which case the (virtual) ISL has VE_Ports at both ends. Figure 11-11 on page 446 shows the connection types between the compute node and the I/O module switches. As shown, if the IBM Flex System Enterprise Chassis I/O module switches were connected they would use VE_Ports at each end. In the IBM Flex System Enterprise Chassis the FCFs will not be connected, the VE_Port connection is used as an illustration example only in Figure 11-11 on page 446.

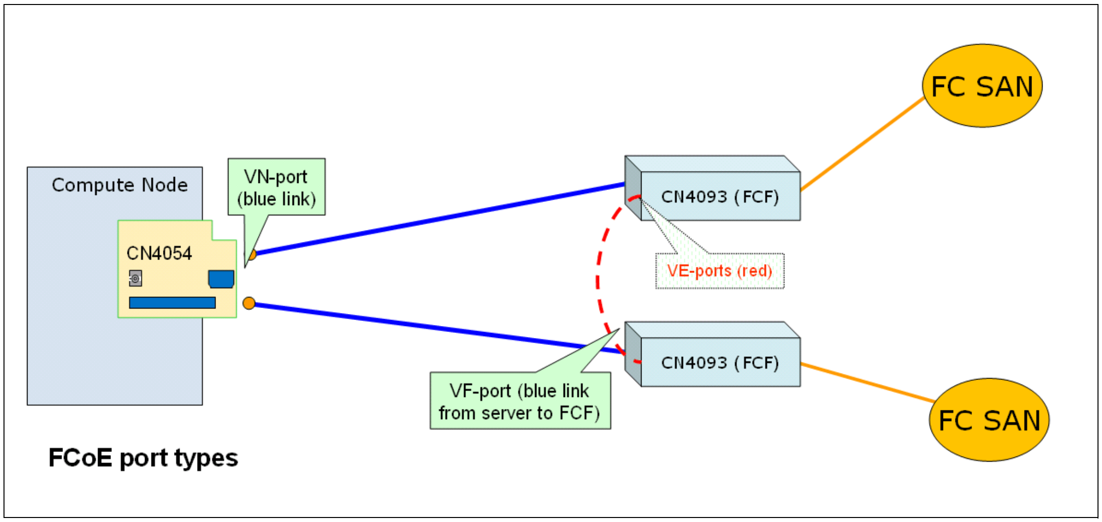


Figure 11-11 Port type between Compute nodes and IBM Flex System Enterprise Chassis switches

FCoE Initialization Protocol and snooping bridges

In a traditional FC network, with point-to-point links between an end device and an FC switch, the end device logs in to the fabric. The device exchanges information with the switch by using well-known addresses over its direct link to the switch. In an FCoE network, with potentially intermediate Ethernet links and possibly switches, these login functions become more complicated. They are handled by a protocol called the *FCoE Initialization Protocol* (FIP).

FIP allows the end device (for example, a compute node with a CNA) to discover FCFs and the VLANs with which to connect to them. Then FIP allows the device to establish those connections, which are the VN_Port to VF_Port virtual links as shown in Figure 11-11.

FIP entails the following high-level steps:

1. The end device or compute node broadcasts a FIP VLAN request to the CN4093 and any other FCF in the Ethernet network.
2. FCFs that have VF_Ports reply with a VLAN Notification frame that lists VLANs that the end device or compute node can use.
3. The compute node discovers FCFs that it can log into by broadcasting a Discovery Solicitation frame in the discovered VLAN.
4. FCFs respond with Discovery Advertisement frames. These frames contain such information as an FCF priority and the identifier of the fabric to which the FCF connects.
5. The end device determines which FCF it wants to connect to for fabric login and sends a FIP Fabric Login (FLOGI) request to the FCF to log in to the fabric.
6. The FCF replies with a FLOGI Accept frame, and then the login is complete. The VN_Port to VF_Port link is now established. The Accept frame also provides a mechanism for the FCF to indicate to the end device the MAC address to use for its VN_Port.

These virtual links can be established over an arbitrary Ethernet network, they must now be given security that is equivalent to the security in a point-to-point FC network. This security is accomplished by having the intermediate DCB-capable CN4093 switch “snoop” the FIP frames that they forward. By using the information that the switch sees during the FIP login sequence, the switch can determine which devices are connected by using a virtual link. Then the switch dynamically creates narrowly tailored access control lists (ACLs) that permit expected FCoE traffic to be exchanged between the appropriate devices and deny all other

undesirable FCoE or FIP traffic. The CN4093 provides FIP snooping function allows the compute node to log in establishing the VN_Port to VF_Port virtual link.

For more information about FIP, see the FC-BB-5 standard at:

http://fcoe.com/09-056v5.pdf

MAC addresses used by end devices

End devices such as the compute nodes (ENodes in the FC-BB-5 standard) use virtual MAC addresses for their VN_Ports. The FC-BB-5 standard allows these MAC addresses to be assigned by the FCF during login or by the ENode. MAC addresses assigned by FCFs are called *Fabric Provided MAC Addresses* (FPMAs). MAC addresses assigned by the end devices are called *Server Provided MAC Addresses* (SPMAs). The CNAs and FCFs today implement only FPMAs hence this is provided by the IBM Flex System Enterprise Chassis I/O module switch such as the IBM Flex System Fabric EN4093 10Gb Scalable Switch.

FCFs, Fabric Mode, and N_Port ID Virtualization

As mentioned previously, a Fibre Channel Forwarder is the FC switching element in an FCoE network. One of the characteristics of an FC switching element is that it joins the FC fabric as a *domain*. This will give it the CN4093 capability of switching data between the compute node using FCoE and an external storage controller attached to external FC SAN fabric. It will also provide connectivity to external FCoE but does not support E-port attachment to switches.

In a mixed FC-FCoE fabric, the FCF also often acts as the conversion device between FC and FCoE. Each FCF that operates in *full-fabric mode* or *switch mode* as an FC switch joins the existing FC fabric as a domain. If the CN4093 is not used in this mode and becomes a gateway device to an external FC or FCoE SAN then *N_Port ID Virtualization* (NPIV) would be employed. NPIV is discussed in detail in “N_Port ID Virtualization” on page 470 where where connections involving NPIV equally apply to FCoE as they do in FC connectivity.

11.4.3 Configuring CN4093 for FCoE Connectivity

This section is specific to IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch I/O module as it has both FC and FCoE capability. There are other I/O module 10Gb Ethernet switches which can be used for FCoE, such as IBM Flex System Fabric EN4093 10Gb Scalable Switch. To Configure FCoE on IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch for connectivity to IBM Flex System V7000 Storage Node, it is necessary to understand the different ports and functions within the switch.

As previously mentioned the CN4093 has FC and FCoE functionality, the physical ports comprise of internal and external types. The internal ports connect the IBM Flex System Enterprise Chassis node bays as shown in Figure 11-1 on page 434. The IBM Flex System V7000 Storage Node uses these internal ports to connect to the IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch. The IBM Omni external ports can be attached via cable to external LAN/SAN network equipment depending whether they are configured for Ethernet or FC mode.

Table 11-2 on page 448 details the different types of ports in the IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch.

Table 11-2 IBM Flex System CN4054 10Gb Virtual Fabric Adapter Ports)

Port Type	Port Name / Range	Description
Ethernet Ports (Internal)	INTA1-INTA14 (ports 1-14), INTB1-INTB14 (15-28), INTC1-INTC14 (29-42)	These standard 10Gb SFP+ Ethernet ports connect internally to compute nodes and storage in the system chassis.
Ethernet Ports (External)	EXT1-EXT2 (ports 43-44)	These standard 10Gb SFP+ Ethernet ports provide external connectors.
High-Capacity Ethernet Ports (External)	EXT3-EXT10 (ports 45-52)	These 40Gb QSFP+ Ethernet ports can be configured as either two 40Gb Ethernet ports (EXT15 and EXT19), or as four 10Gb Ethernet ports (EXT15-EXT18, EXT19-EXT22)
IBM Omni Ports (External)	EXT11-EXT22 (ports 53-64)	These 10Gb SFP+ hybrid ports can be configured to operate either in Ethernet mode (the default) or in Fibre Channel mode for direct connection to Fibre Channel devices.

The Omni ports are all by default in Ethernet mode and can carry Ethernet traffic as well as FCoE. These ports can be configured as FC ports and attach to external storage controllers or servers. They must be configured in pairs for either Ethernet or FC mode, for example, port EXT11 and EXT12 must both be configured to be in the same mode and so are called paired ports.

Fibre Channel VLANs

The ports that will be used to connect using FCoE must be isolated into a separate VLAN on the IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch, the VLAN when defined must have a VLAN number and the following components:

- Port Membership - named ports, as detailed in Table 11-2 the VLAN must include at least one FC defined ports (paired FC Omni ports can reside in a separate FC VLAN)
- Switch Role - this is either full switch fabric or NPV mode

The switch role mode for that VLAN will determine whether it has the switching element thus FCF capability, or has to pass all data to an external SAN switch for redirection thus NPV capability. For a compute node to connect to IBM Flex System V7000 Storage Node and access storage the VLAN must have FCF enabled. With this, all storage data will remain within the IBM Flex System Enterprise Chassis and not have to rely on an external SAN switch for its switching or redirection.

Figure 11-12 on page 449 shows VLAN 1002 which has been created and includes external ports EXT11 and EXT12 along with internal ports from the IBM Flex System V7000 Storage Node canisters. The storage node occupies node ports 1-4 in the IBM Flex System Enterprise Chassis so ports INTA1-4 are available for this VLAN, of which INTA3+4 have been selected. The port from Compute Node 11 (INTA11) has also been included in the Fibre Channel VLAN.

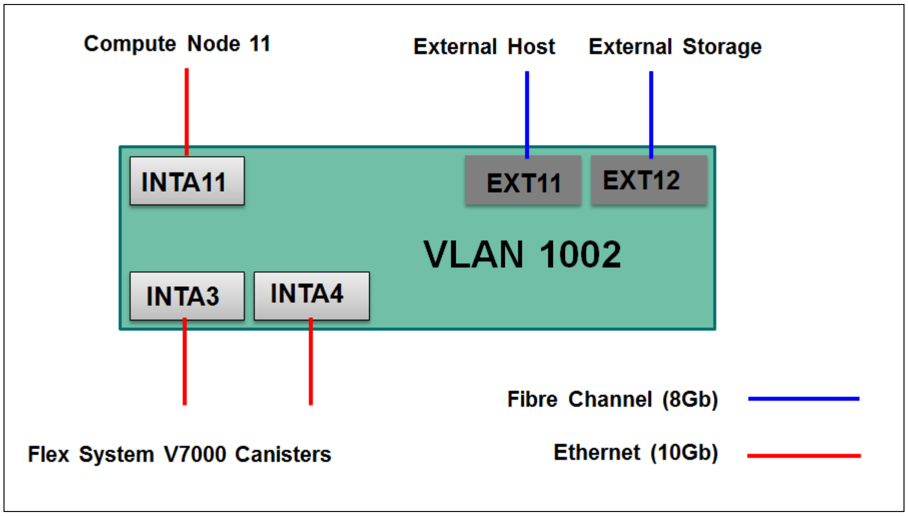


Figure 11-12 Internal and External port inclusion for FCoE VLAN

With this VLAN created, FCoE zones can be configured to map Compute Node 11 to IBM Flex System V7000 Storage Node via internal ports INTA3+4, and/or to external storage via port EXT12. The connectivity between Compute Node 11 and IBM Flex System V7000 Storage Node is FCoE as shown in Figure 11-12 but any connection to external storage via port EXT12 is using fibre channel. The IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch with this VLAN configured and using FCF provides an example of FCoE gateway for bridging FCoE and Fibre Channel networks. This is where Compute Node 11 using FCoE connectivity can attach to external storage which is FC attached to the CN4093.

Administration Interface for CN4093

There are several methods of accessing CN4093 in order to configure, view or make changes:

1. Using a Telnet/SSH connection via the chassis management module
2. Using a Telnet/SSH connection over the network
3. Using the Browser-Based Interface (BBI) over the network
4. Using a serial connection via the serial port on the CN4093

The Telnet/SSH connection accesses two types of Command Line Interpreter (CLI), one is a text menu based CLI, the second is based on the international standard CLI (ISCLI). In this section we use the ISCLI to display and enter commands on the IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch. For details on configuration and access for configuration see << Author Comment - link to CN4093 install manual and CLI reference manual (not yet available)>>

Configuring for Fibre Channel Forwarding

In this section we will create the VLAN as shown in Figure 11-12 before creating zones and allowing access from Compute Node 11 to the IBM Flex System V7000 Storage Node located in the first 4 bays of the IBM Flex System Enterprise Chassis. A summary of the steps are:

1. Enable CEE and FIPS
2. Convert the omni-ports EXT11 and EXT12 (53,54) to Fibre Channel ports
3. Create FCoE VLAN
4. Assign ports to FCoE VLAN
5. Enable FCF (must have at least one FC-Port to enable FCF)

It is necessary to complete the above steps in the correct order for the configuration to be successful. In Example 11-1 the ISCLI commands display shows the Omni ports EXT11-12 are changed from their ethernet default to Fibre Channel after the Converged Enhanced Ethernet (CEE) and FCoE Initialization Protocol (FIP) snooping is enabled. The FCoE VLAN is then created and then the ports are assigned to the VLAN

Example 11-1 Configuring ports, creating VLAN and port assigning

```
Router(config)#cee enable------------------------------------------> CEE Enabled
Router(config)#cfcoe fips enable-----------------------------------> FIPS Enabled
Router(config)#system port EXT11-EXT12 type fc ---------------> FC Mode Enabled
Router(config)#
Oct 26 14:22:12 fd8c:215d:178e:c0de:3640:b5ff:fe5c:f5ef NOTICE  lldp: LLDP TX & RX
are disabled on port EXT11

Oct 26 14:22:12 fd8c:215d:178e:c0de:3640:b5ff:fe5c:f5ef NOTICE  lldp: LLDP TX & RX
are disabled on port EXT12

Router(config)#vlan 1002

VLAN 1002 is created------------------------------------------------>VLAN Created
Router(config-vlan)#member INTA3-INTA4,INTA11-------------> Ports added to VLAN
Port INTA3 is an UNTAGGED port and its PVID is changed from 1 to 1002
Port INTA4 is an UNTAGGED port and its PVID is changed from 1 to 1002
Port INTA11 is an UNTAGGED port and its PVID is changed from 1 to 1002
Router(config-vlan)#member EXT11-EXT12--------------------> Ports added to VLAN
Router(config-vlan)
```

Example 11-2 shows an abbreviated version of the ISCLI command output for ***show fc-internals*** command, this shows that the Omni ports EXT11 and EXT12 have been FC enabled and are assigned to VLAN 1002.

Example 11-2 Checking Omni ports are correctly configured for FC

```
Alias   Port   FC   VLAN   VLAN-Mode   EACL   IPG
-----   ----   --   ----   ---------   ----   ---
 EXT11      53      Y    1002        FCF       Y       Y
 EXT12      54      Y    1002        FCF       Y       Y
 EXT13      55      N       1        N/A       N       N
 EXT14      56      N       1        N/A       N       N
 EXT15      57      N       1        N/A       N       N
 EXT16      58      N       1        N/A       N       N
```

Example 11-3 shows all ports have been successfully added to VLAN 1002 with VLAN enabled.

Example 11-3 Display VLAN and ports

```
Router(config-vlan)#show vlan
VLAN                Name                Status MGT           Ports
----  --------------------------------  ------ --- -------------------------
1     Default VLAN                      ena    dis   INTA1 INTA2 INTA5-INTA10
                                                     INTA12-INTA14 EXT1 EXT2
                                                     EXT11-EXT16
1002  VLAN 1002                         ena    dis   INTA3 INTA4 INTA11 EXT11
                                                     EXT12
4095  Mgmt VLAN                         ena    ena   EXTM MGT1
```

```
Router(config-vlan)#
```

The next step is to enable Fibre Channel Forwarding (FCF) where Example 11-4 shows the ISCLI command ***fcf enable*** run where on completion FCoE connections are established.

Example 11-4 Enable Fibre Channel Forwarding

```
Router(config-vlan)#fcf enable------------------------------------>FCF Enabled
Router(config-vlan)#
Oct 26 14:26:04 fd8c:215d:178e:c0de:3640:b5ff:fe5c:f5ef NOTICE  fcoe: FCOE
connection between VN_PORT 0e:fc:00:01:0d:00 and FCF 34:40:b5:5c:ee:03 has been
established.

Oct 26 14:26:04 fd8c:215d:178e:c0de:3640:b5ff:fe5c:f5ef NOTICE  fcoe: FCOE
connection between VN_PORT 0e:fc:00:01:0c:00 and FCF 34:40:b5:5c:ee:02 has been
established.
```

This completes the FCF component. To check our configuration is correct we can look at the FCoE database where the port world wide names (PWWN) can be seen and so used for zoning. Example 11-5 shows the output of the ISCLI ***show fcoe database*** command where connections are established between the IBM Flex System V7000 Storage Node on ports INTA3+4 and the Compute Node 11 in bay 11 has FCoE configured and also established a connection from port INTA11.

Example 11-5 Checking the FCoE database

```
Router(config)#show fcoe database
-----------------------------------------------------------------------
 VLAN  FCID                 WWN                      MAC          Port
-----------------------------------------------------------------------
 1002  010c00     50:05:07:68:05:04:02:50     0e:fc:00:01:0c:00   INTA3
 1002  010d00     50:05:07:68:05:04:02:51     0e:fc:00:01:0d:00   INTA4
 1002  010c01     10:00:5c:f3:fc:6e:30:01     0e:fc:00:01:0c:01   INTA11

 Total number of entries = 3
-----------------------------------------------------------------------
Router(config)#
```

We can also confirm connectivity from IBM Flex System V7000 Storage Node by reviewing the ‘System Details’ option from the administration GUI. Figure 11-13 on page 452 shows Canister 1 where 10Gb Ethernet port is active detailing the Port World Wide Name (PWWN) or WWPN in screen shots below.

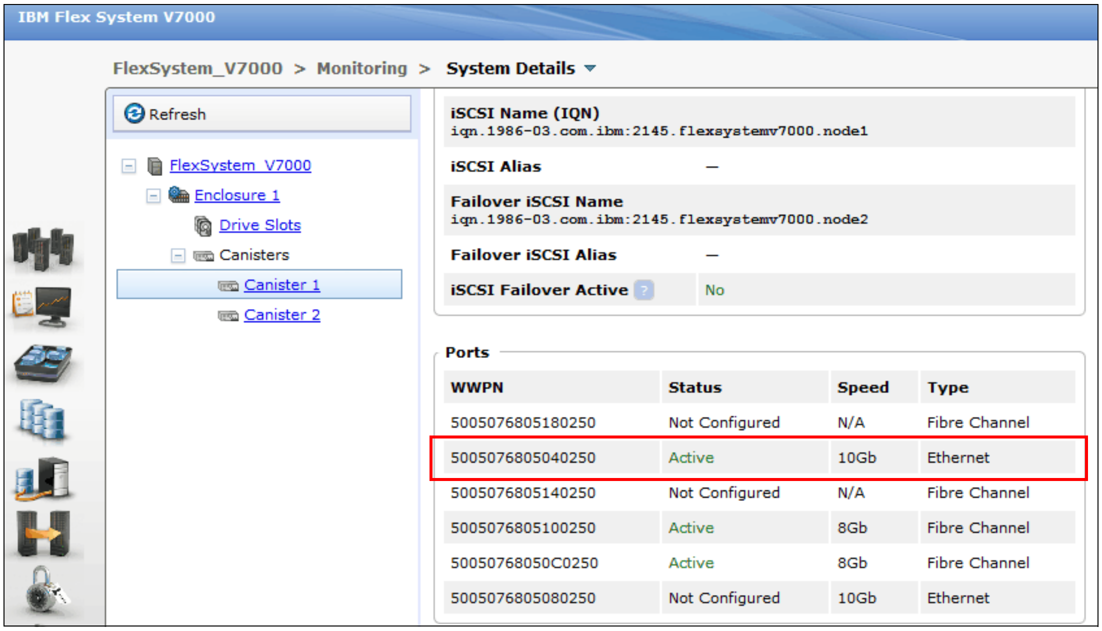


Figure 11-13 Active 10Gb Ethernet on Canister 1

Figure 11-14 shows the active Ethernet port for Canister 2 of the IBM Flex System V7000 Storage Node. By comparing the canister PWWNs with Example 11-5 on page 451 you can see that Canister 1 uses port INTA3 and Canister 2 uses INTA4.

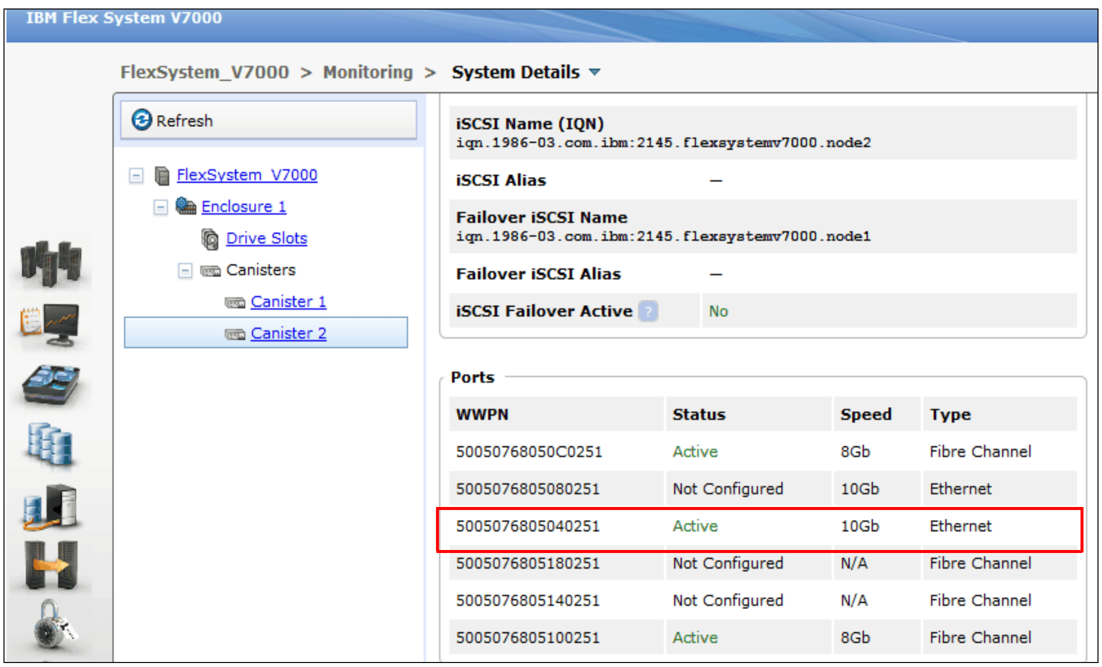


Figure 11-14 Active 10Gb Ethernet on Canister 2

Create FCoE Zoning

By creating a zone which will include PWWNs of the host - Compute Node 11 and the storage controller - IBM Flex System V7000 Storage Node the two end points will be able to connect and storage can be accessed by the O/S platform residing on Compute Node 11. The following zoning steps include:

- Create Zone
- Create Zoneset
- Activate Zoneset

This process is very much the same as is used on FC zoning. Example 11-6 shows from the ISCLI creating a zones and populating it with PWWNs from Compute Node 11 and each canister of the IBM Flex System V7000 Storage Node.

Example 11-6 Creating Zone and Zoneset

```
Router(config)#zone name FSV7K_N11---------------------------------> Create Zone
Router(config-zone)#member pwwn 50:05:07:68:05:04:02:50
Router(config-zone)#member pwwn 50:05:07:68:05:04:02:51
Router(config-zone)#member pwwn 10:00:5c:f3:fc:6e:30:01--------> Add PWWN members
Router(config-zone)#show zone
  zone name FSV7K_N11
      pwwn 50:05:07:68:05:04:02:50
      pwwn 50:05:07:68:05:04:02:51
      pwwn 10:00:5c:f3:fc:6e:30:01
Router(config-zone)#zoneset name CN4093_1------------------------> Create Zoneset
Router(config-zoneset)#member FSV7K_N11---------------------> Add Zone to Zoneset
Router(config-zoneset)#show zoneset-------------------------------> Check Zoneset
zoneset name CN4093_1
  zone name FSV7K_N11
      pwwn 50:05:07:68:05:04:02:50
      pwwn 50:05:07:68:05:04:02:51
      pwwn 10:00:5c:f3:fc:6e:30:01
zoneset name FCOE_Zoneset1
```

Example 11-7 shows from the ISCLI activating the zoneset and then checking to ensure the configuration and active zoneset are correct.

Example 11-7 Activating and checking zoneset

```
Router(config-zoneset)#zoneset activate name CN4093_1----------> Activate Zoneset
Router(config)#show zoneset active-----------------------------> Check Activation
Active Zoneset CN4093_1 has 1 zones

zoneset name CN4093_1
  zone name FSV7K_N11
      pwwn 50:05:07:68:05:04:02:50
      pwwn 50:05:07:68:05:04:02:51
      pwwn 10:00:5c:f3:fc:6e:30:01

Default-Zone Deny
Router(config)#
```

Once this is successfully completed the PWWN should be visible from the IBM Flex System V7000 Storage Node where a host can be created and storage allocated.

Figure 11-15 on page 454 shows the PWWN from the FCoE attached Compute Node 11 is seen from the administrative GUI of the IBM Flex System V7000 Storage Node. It can be used to define a host and host port so that volumes can be assigned.

It is important to remember that this entire process has to be repeated on the IBM Flex System CN4054 10Gb Virtual Fabric Adapter residing in I/O Module port 2 in the IBM Flex

System Enterprise Chassis where a second host port can be added to the host name. This will eliminate one of the switches as being a point of failure. The same VLAN number must be used, in the above example we used VLAN 1002.

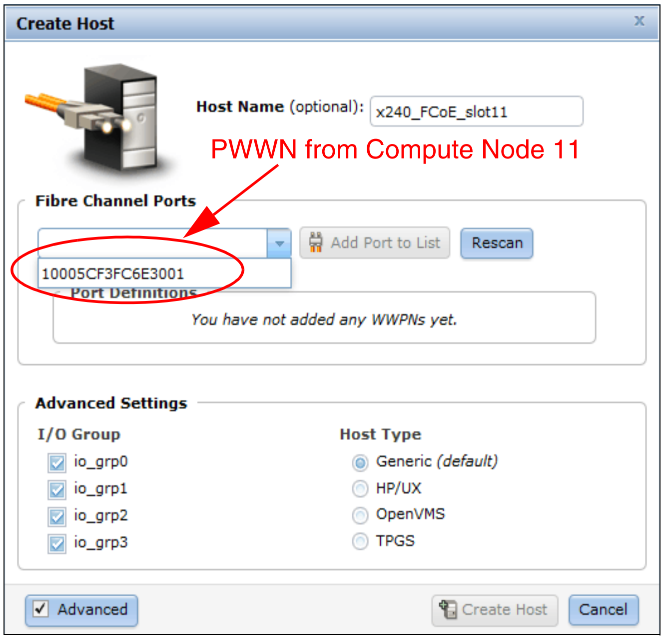


Figure 11-15 Host and port to be defined using FCoE PWWN

When both IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch in the IBM Flex System Enterprise Chassis are configured the second 10Gb Ethernet port in each canister becomes active completing the configuration. Figure 11-16 shows Canister 1 with both 10Gb Ethernet ports active completing connections to both I/O module switches.

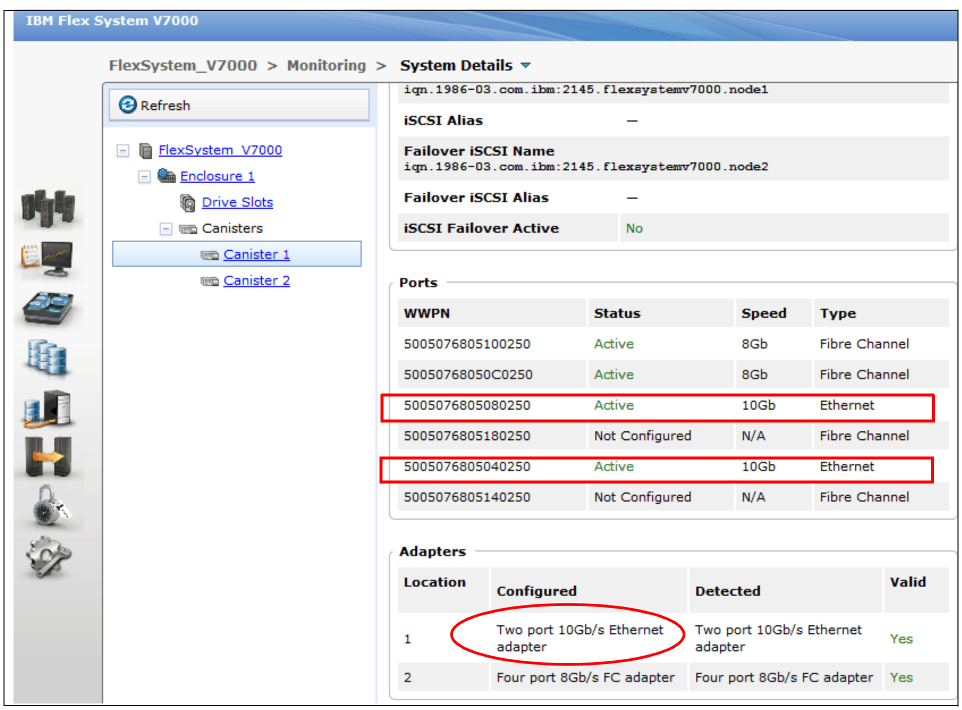


Figure 11-16 Canister 1 with both ethernet ports active

11.5 Fibre Channel Connectivity

Fibre Channel (FC) is an open, technical standard for networking which incorporates the “channel transport” characteristics of an I/O bus, with the flexible connectivity and distance characteristics of a traditional network. Due to its channel-like qualities, hosts and applications see storage devices attached to the SAN as though they are locally attached storage. With its network characteristics FC can support multiple protocols and a broad range of devices, and it can be managed as a network.

Like other networks, information is sent in structured packets or frames, and data is serialized before transmission. But, unlike other networks, the Fibre Channel architecture includes a significant amount of hardware processing to deliver high performance. Fibre Channel uses a serial data transport scheme, similar to other computer networks, streaming packets, (frames) of bits one behind the other in a single data line to achieve high data rates. Serial transfer by its very nature does not suffer from the problem of skew, so speed and distance is not restricted as with parallel data transfers as shown in Figure 11-17. This was a distinct advantage over parallel SCSI data buses which, along with other SCSI bus limitations confined the bus to limited distance or internally attached server storage.

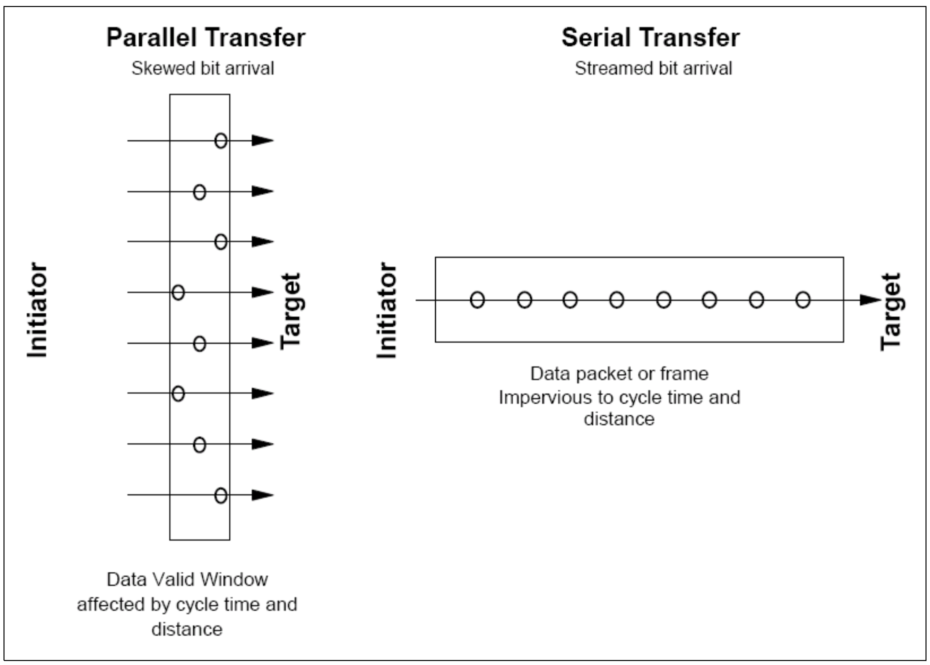


Figure 11-17 Parallel and serial data transfer types

Serial transfer enables simpler cabling and connectors, and also routing of information through switched networks. Fibre Channel can operate over longer distances, both natively and by implementing cascading, and longer distances with the introduction of repeaters. Just as LANs can be interlinked in WANs by using high speed gateways, so can campus SANs be interlinked to build enterprise wide SANs.

Whatever the topology, information is sent between two nodes, which are the source (transmitter or initiator) and destination (receiver or target). A node is a device, such as a server (compute node), or peripheral device, such as IBM Flex System V7000 Storage Node or tape drive. Frames of information are passed between nodes, and the structure of the frame is defined by a protocol. Logically, a source and target node must utilize the same protocol, but each node may support several different protocols or data types. In the IBM Flex

System Enterprise Chassis the compute nodes are the initiators with IBM Flex System V7000 Storage Node being the target.

Fibre Channel architecture is extremely flexible in its potential application. Fibre Channel transport layers are protocol independent, enabling the transmission of multiple protocols.

Using a credit based flow control methodology, Fibre Channel is able to deliver data as fast as the destination device buffer is able to receive it. And low transmission overheads enable high sustained utilization rates without loss of data. Therefore, Fibre Channel combines the best characteristics of traditional I/O channels with those of computer networks such as within the IBM Flex System Enterprise Chassis.

Fibre Channel Layers

Fibre Channel (FC) is broken up into a series of five layers. The concept of layers, starting with the ISO/OSI seven-layer model, allows the development of one layer to remain independent of the adjacent layers. Although, FC contains five layers, those layers follow the general principles stated in the ISO/OSI model.

The five layers can be categorized into two:

- Physical and signaling layer
- Upper layer

Fibre Channel is a layered protocol. as shown in Figure 11-18, which can be compared to Figure 11-8 on page 443 where the FC layer can be seen to be encapsulated within the FCoE stack.

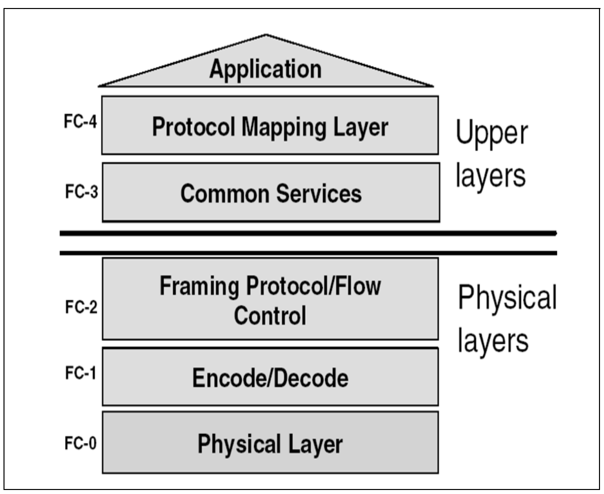


Figure 11-18 Upper and physical FC Layers

Physical and signaling layers

The physical and signaling layers include the three lowest layers: FC-0, FC-1,and FC-2.

- Physical interface and media: FC-0
  – The lowest layer, FC-0, defines the physical link in the system, including the cabling, connectors, and electrical parameters for the system at a wide range of data rates. This level is designed for maximum flexibility, and allows the use of a large number of technologies to match the needs of the configuration.
- Transmission protocol: FC-1
  – The second layer, FC-1 provides the methods for adaptive 8B/10B encoding to bind the maximum length of the code, maintain DC-balance, and provide word alignment. This

layer is used to integrate the data with the clock information required by serial transmission technologies.

- Framing and signaling protocol: FC-2
  - Reliable communications result from Fibre Channel's FC-2 framing and signaling protocol. FC-2 specifies a data transport mechanism that is independent of upper layer protocols.

Upper layers

The Upper layer includes two layers: FC-3 and FC-4.

- Common services: FC-3
  - FC-3 defines functions that span multiple ports on a single-node or fabric. Functions that are currently supported include:
    - Hunt Groups – A a set of associated N_Ports attached to a single node. This set is assigned an alias identifier that allows any frames containing the alias to be routed to any available N_Port within the set. This decreases latency in waiting for an N_Port to become available.
    - Striping – Striping is used to multiply bandwidth, using multiple N_Ports in parallel to transmit a single information unit across multiple links.
    - Multicast – Multicast delivers a single transmission to multiple destination ports. This includes the ability to broadcast to all nodes or a subset of nodes.
- Upper layer protocol mapping (ULP): FC-4
  - The highest layer, FC-4, provides the application-specific protocols. Fibre Channel is equally adept at transporting both network and channel information and allows both protocol types to be concurrently transported over the same physical interface.

A channel protocol example is Fibre Channel Protocol (FCP). This is used to transfer SCSI commands and data over Fibre Channel and is commonly used in FC SANs.

Fibre Channel combines the best characteristics of traditional I/O channels with those of computer networks:

- High performance for large data transfers by using simple transport protocols and extensive hardware assists
- Serial data transmission
- A physical interface with a low error rate definition
- Reliable transmission of data with the ability to guarantee or confirm error free delivery of the data
- Packaging data in packets (frames in Fibre Channel terminology)
- Flexibility in terms of the types of information which can be transported in frames (such as data, video and audio)
- Use of existing device oriented command sets, such as SCSI and FCP
- A vast expansion in the number of devices which can be addressed when compared to I/O interfaces — a theoretical maximum of more than 15 million ports

It is this high degree of flexibility, availability and scalability; the combination of multiple protocols at high speeds over long distances; and the broad acceptance of the Fibre Channel standards by vendors throughout the IT industry, which makes the Fibre Channel architecture ideal for the development of enterprise SANs.

For further information on Fibre Channel layers and SAN networking *Introduction to Storage Area Networks and System Networking*, SG24-5470

11.5.1 Fibre Channel Topologies

Fibre Channel provides three distinct and one hybrid interconnection topologies. By having more than one interconnection option available, a particular application can choose the topology that is best suited to its requirements.

The three Fibre Channel topologies are:

- Point-to-point, the point-to-point topology consists of a single connection between two nodes. All the bandwidth is dedicated to these two nodes.
- Arbitrated loop, Loop, in the loop topology, the bandwidth is shared between all the nodes connected to the loop. The loop can be wired node-to-node; however, if a node fails or is not powered on, the loop is out of operation. This is overcome by using a hub. A hub opens the loop when a new node is connected, and closes it when a node disconnects. Although more popular a number of years ago when FC speeds were slower, this topology is rarely used now.
- Switched or fabric, a switch allows multiple concurrent connections between nodes. There can be two types of switches: circuit switches and frame switches. Circuit switches establish a dedicated connection between two nodes, where as frame switches route frames between nodes and establish the connection only when needed. This is also known as switched fabric.

The point-to-point topology is used within the IBM Flex System Enterprise Chassis switch between the compute node and the I/O Module FC switch. Also between the switch and IBM Flex System V7000 Storage Node. The switch or fabric topology is used between the IBM Flex System Enterprise Chassis switches and external FC SAN.

Figure 11-19 on page 459 shows the point to point topology used in the FC connectivity within the IBM Flex System Enterprise Chassis. In this example there is a switched topology between the I/O Module switch FC3171 and the external FC switch.

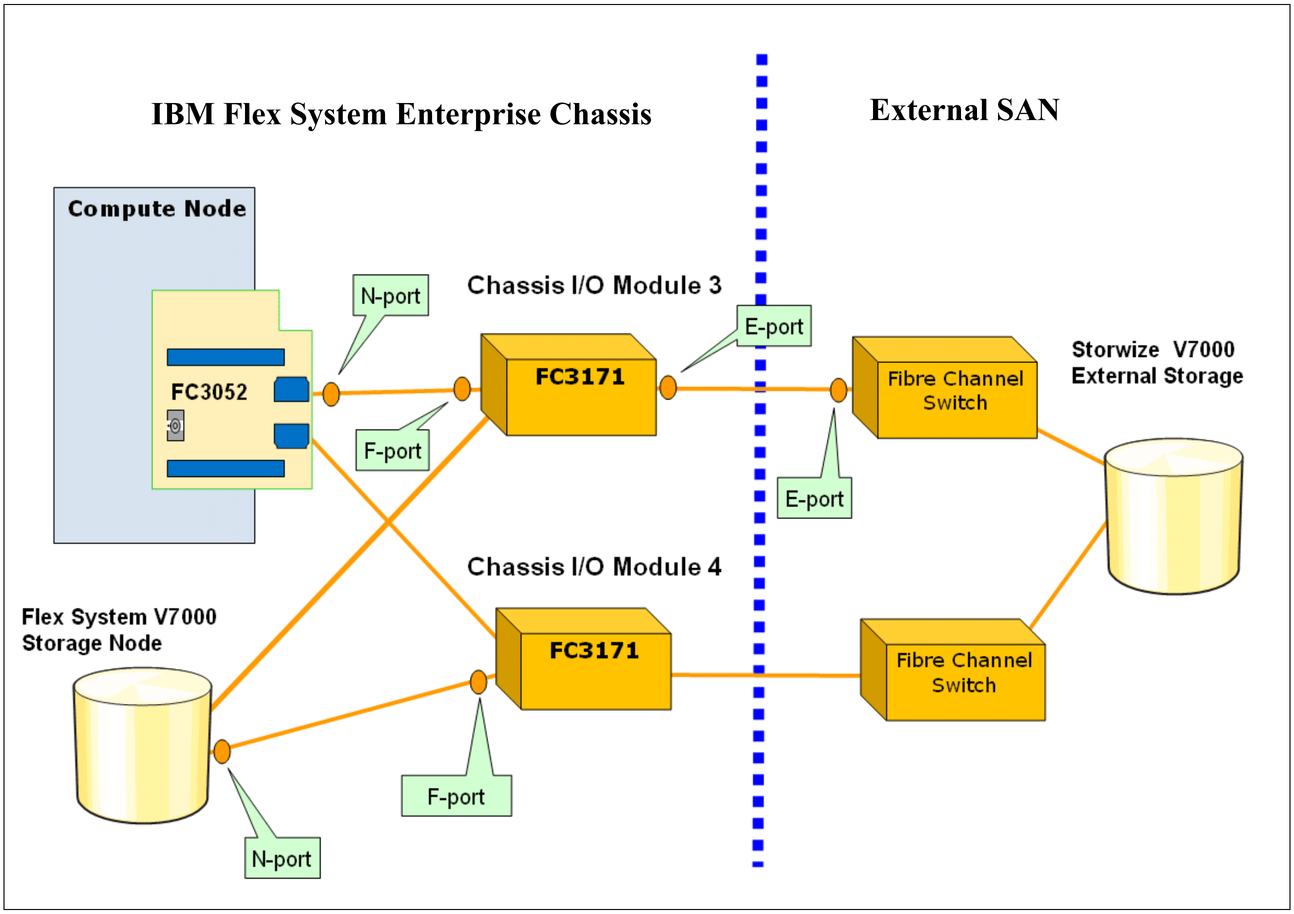


Figure 11-19 FC point to point topology within the IBM Flex System Enterprise Chassis

11.5.2 FC addressing and port types

In a Fibre Channel environment the unique identity of participants are maintained through a hierarchy of fixed names and assigned addresses identifiers. In Fibre Channel terminology, a communicating device is a node. Each node has a fixed 64-bit Node_name assigned by the manufacturer. The node name will be unique if the manufacturer has registered a range of addresses with the IEEE, and so is normally referred to as a World-Wide Name. An N_Port within a parent (WWN) node is also assigned a unique 64-bit Port_Name, which aids the accessibility of the port and is known as the Port World-Wide Name (PWWN).

In FC SAN environments, the FC switch is responsible for assigning a 24-bit address to each device as it logs on. Allowing the topology to manage the assignment of addresses has the advantage that control of the addresses is now performed by the entity that is responsible for the routing of information. This means that address assignments can be made in a manner that results in the most efficient routing of frames within that topology.

There are six standard kinds of ports, those used with IBM Flex System V7000 Storage Node are listed below:

– Node ports (N_Port), these ports are found in Fibre Channel nodes, which are defined to be the source or destination of information units (IU). I/O devices and host systems interconnected in point-to-point or switched topologies use N_Ports for their connection. N_Ports can only attach to other N_Ports or to F_Ports.

- Fabric ports (F_Port), these ports are found in Fibre Channel switched fabrics. They are not the source or destination of IU's, but instead function only as a "middle-man" to relay the IUs from the sender to the receiver. F_Ports can only be attached to N_Ports.
- Expansion ports (E_Port), these ports are found in Fibre Channel switched fabrics and are used to interconnect the individual switch or routing elements. They are not the source or destination of IUs, but instead function like the F_Ports to relay the IUs from one switch or routing elements to another. E_Ports can only attach to other E_Ports.

The E-Port is not strictly used within the chassis but as shown in the example in Figure 11-19 on page 459 it may be used to connect to external SAN fabric switches. Thus allowing integration of the SAN fabric at switching level for IBM Flex System V7000 Storage Node to access and manage external storage controllers.

11.5.3 Zoning a Compute Node for Storage Allocation

Connectivity over Fibre Channel between the compute nodes and IBM Flex System V7000 Storage Node is achieved through the adapter in the computer node and Fibre Channel switches in the IBM Flex System Enterprise Chassis I/O Module slots. In this section we will look at the connectivity between the FC switches in IBM Flex System Enterprise Chassis I/O Module slots 3 and 4 and configure zones to an IBM Flex System x240 Compute Node. The FC switches used are the IBM Flex System FC3171 8Gb SAN Switch models.

Although the SAN connecting components are within the IBM Flex System Enterprise Chassis with both FC switches in I/O Module slots, the result is a basic and small pair of SAN fabrics. Each have the capability of connecting to external SAN fabrics as they have external ports (label 2 in Figure 11-21 on page 461). Also, within the chassis there will be a requirement to segregate FC initiators from each other and maintain separation of compute nodes. For this reason SAN zoning rules must be applied and the IBM Flex System FC3171 8Gb SAN Switch in each slot must be zoned.

In the example below we show how to connect to, log in and start the zoning configuration process on one FC3171 8Gb SAN Switch. The procedure has to be repeated on the second switch in the IBM Flex System Enterprise Chassis. To access the switch we used browser option for ease of use, although there is CLI which can be accessed via a SSH session. The browser we used below was Firefox which used Java to create the configuration windows.

The switches IP address can be ascertained from the CMM as shown in 3.2.4, "Accessing I/O Modules using CMM" on page 94.

From your chosen browser enter the IP address of the switch:

```
https://xxx.xxx.xxx.xxx
```

Figure 11-20 on page 461 shows the login prompt from the Java window, the default user is 'USERID' with the password 'PASSW0RD' (the '0' in the password is a 'zero').

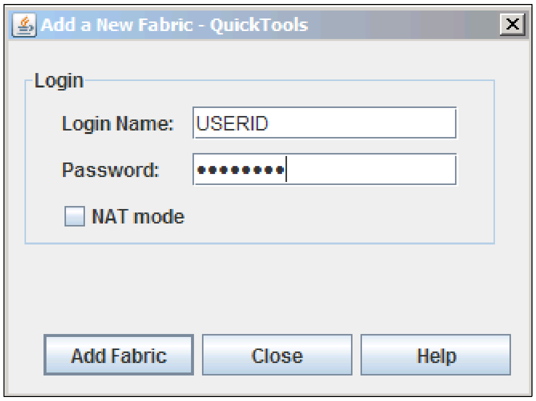


Figure 11-20 Login prompt for FC switch access

Figure 11-21shows the first screen when logged in initially, the bay ports and adapters are displayed with ports which have logged into the switch. The IBM Flex System V7000 Storage Node occupies node bays 1-4. The positions can be verified with the CMM chassis display.

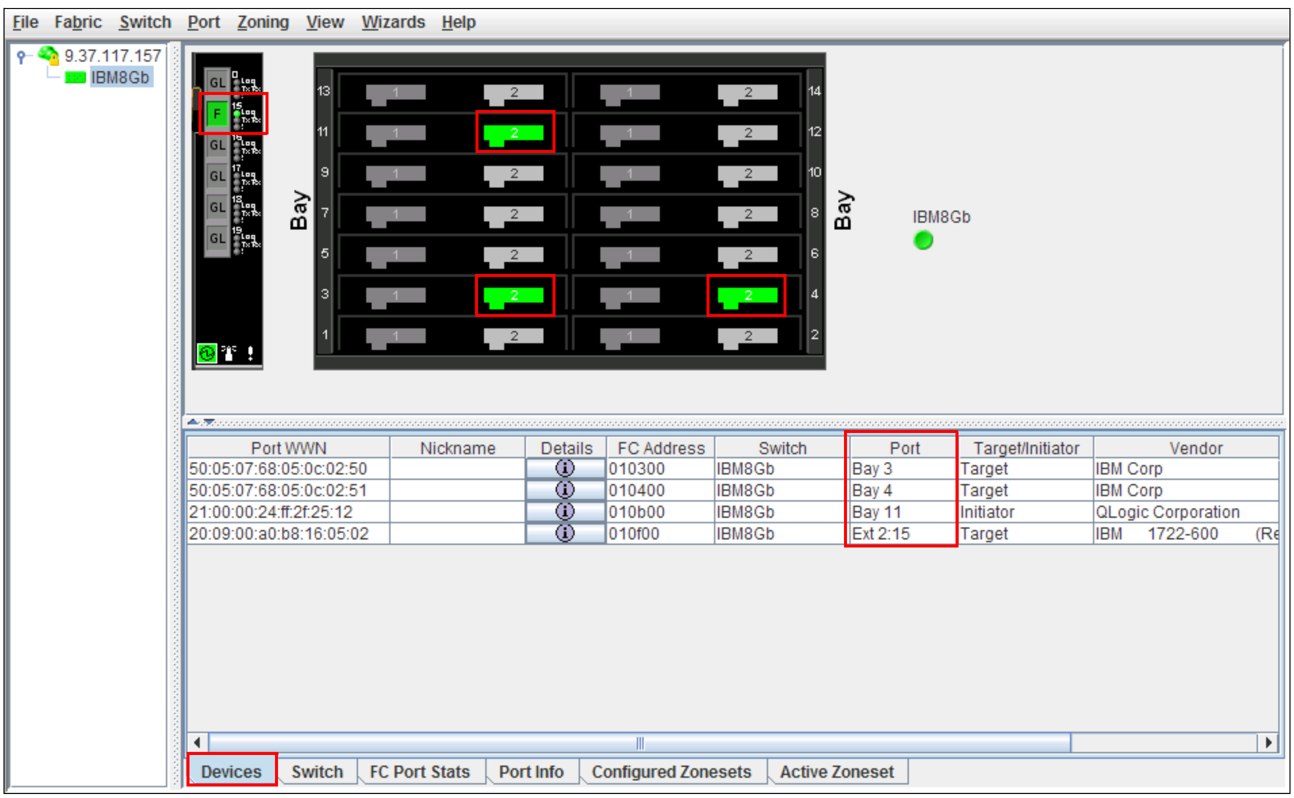


Figure 11-21 Initial FC3171 8Gb SAN Switch configuration

Figure 11-21 has the following indicated:

- External port connection from external storage controller (DS4000®)
- x240 compute node FC adapter in Bay 11 adapter slot 2
- IBM Flex System V7000 Storage Node Network adapter 2 from each canister
- Compute node bays listed with WWN ids in first row
- Indicates view is shown from the ‘Devices’ tab

This switch has not been used before so in this example we will be:

1. creating a zoneset
2. creating aliases
3. creating a zone
4. adding aliases to a zone
5. adding a zone to zoneset
6. activating the zoneset

Figure 11-22 shows from the main menu how to access and configure zoning by clicking on ‘Zoning’ and select ‘Edit Zoning ...’ option.

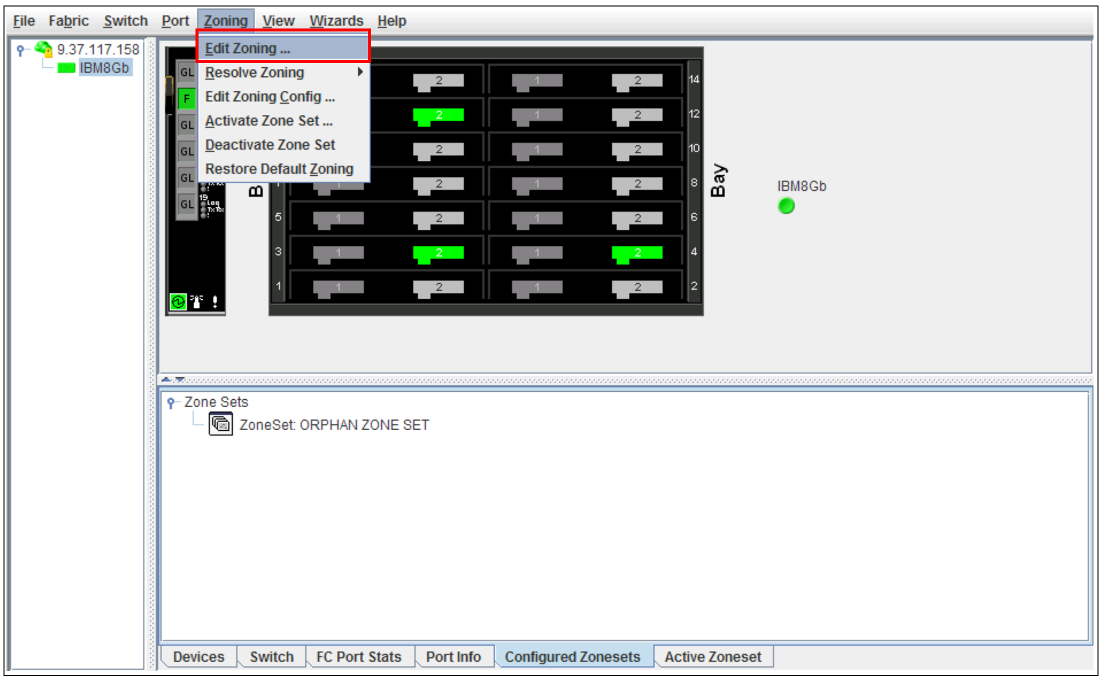


Figure 11-22 Access the zoning configuration window

Creating a Zoneset

From this option the screen in Figure 11-23 on page 463 is displayed where:

1. Select option and click to create zoneset
2. Zoneset creation box will appear where you can type the zoneset name

When you have named the zoneset and clicked the ‘OK’ box, the new zoneset will be created and displayed in the left hand field of the main window.

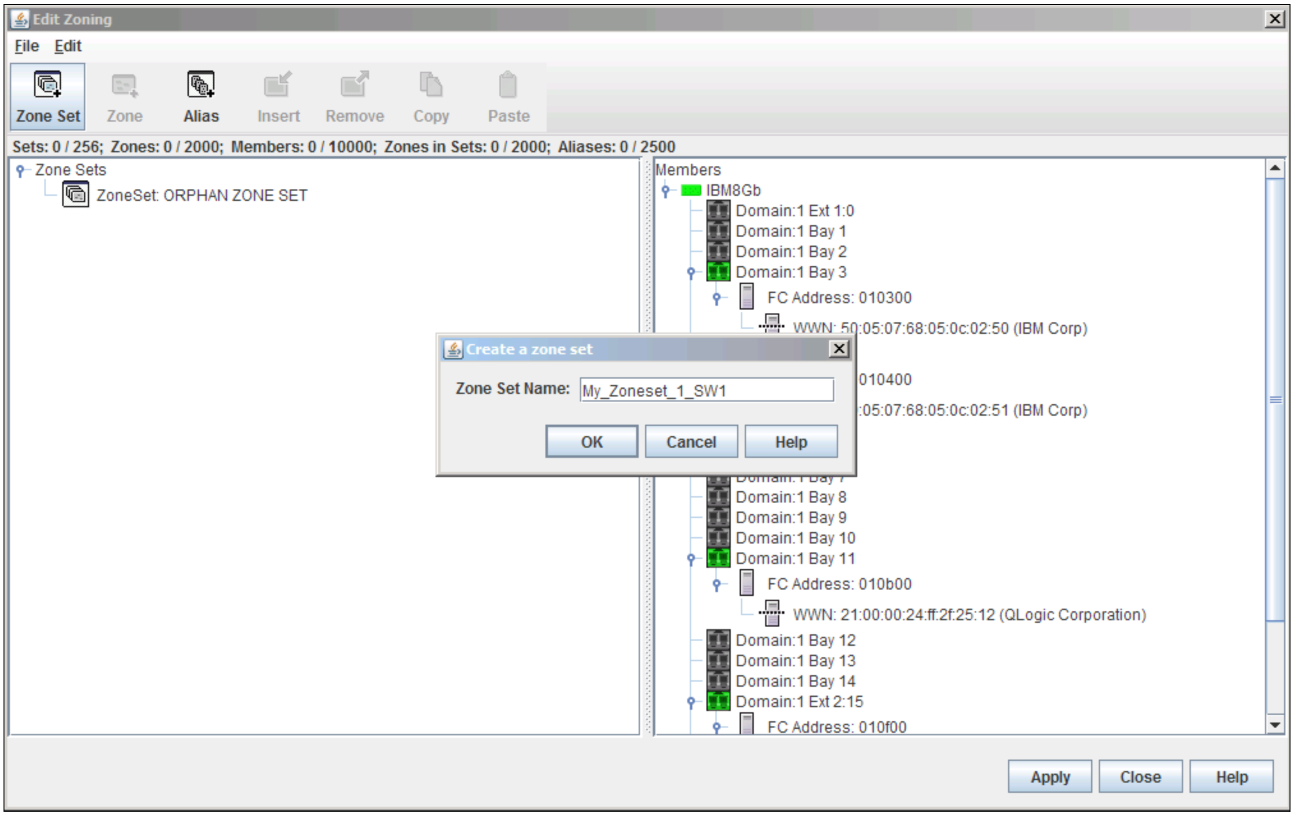


Figure 11-23 Creating a zoneset

Creating an Alias

The next task is to create aliases for the WWN (shown in the right hand display in Figure 11-23), to do this from the main zoning screen, click the ‘Alias’ box and enter the alias name in the ‘Create Alias’ window as shown in Figure 11-24 on page 464 and click the ‘OK’ box when completed.

Note: When creating or changing zonesets, aliases and zones, all changes will not be made to the switch until the ‘Apply’ box has been clicked with the mouse at the bottom right of the zoning window as shown in Figure 11-23

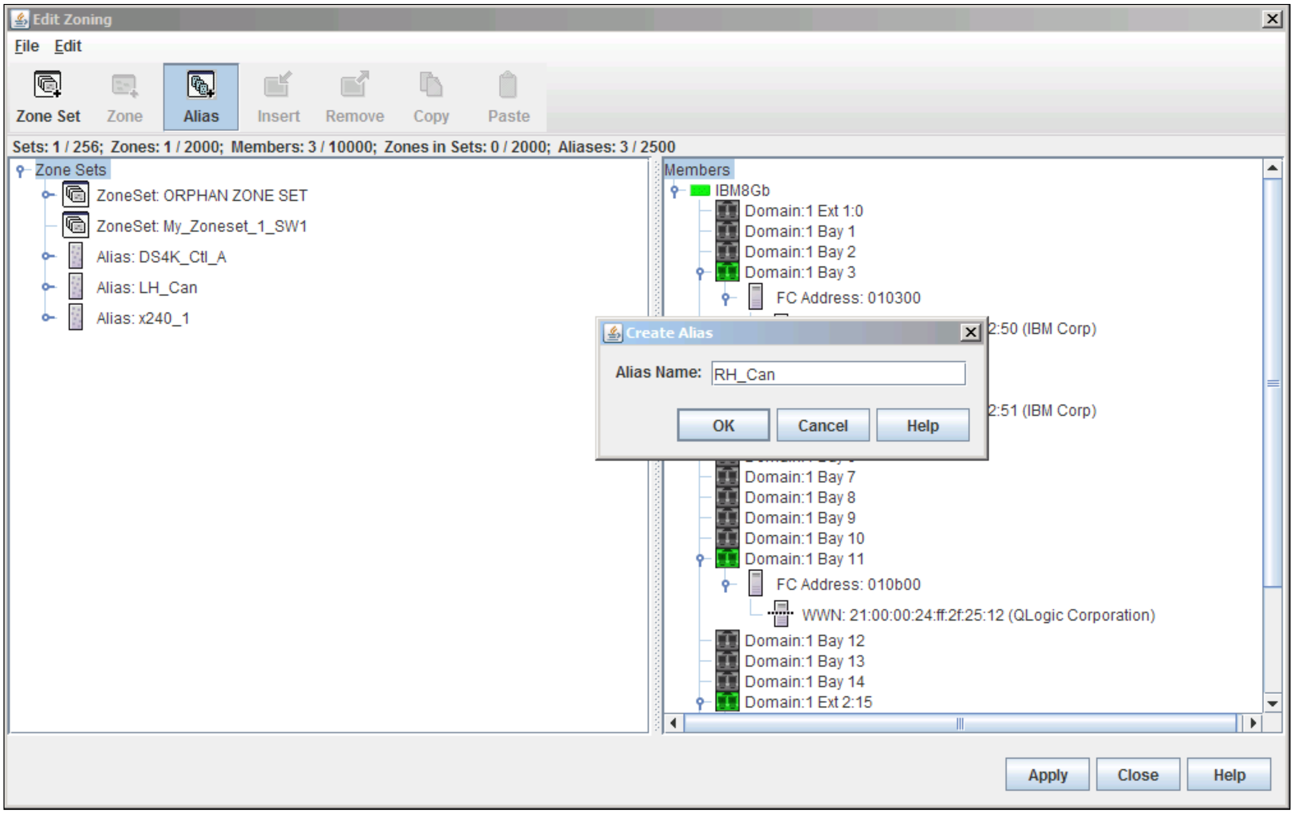


Figure 11-24 Creating an alias name

A WWN has to be added to the alias just created. In Figure 11-25 on page 465 with the mouse, highlight the alias you wish to use and the WWN that you want to add. In this example we have created the alias ‘RH_Can’ and now wish to add the WWN as seen from the right hand column. The WWN in Bay 4 is from the right hand canister so have selected this to add.

When both alias name and WWN name are highlighted with the mouse, click on the ‘Insert’ box as shown in Figure 11-25 on page 465.

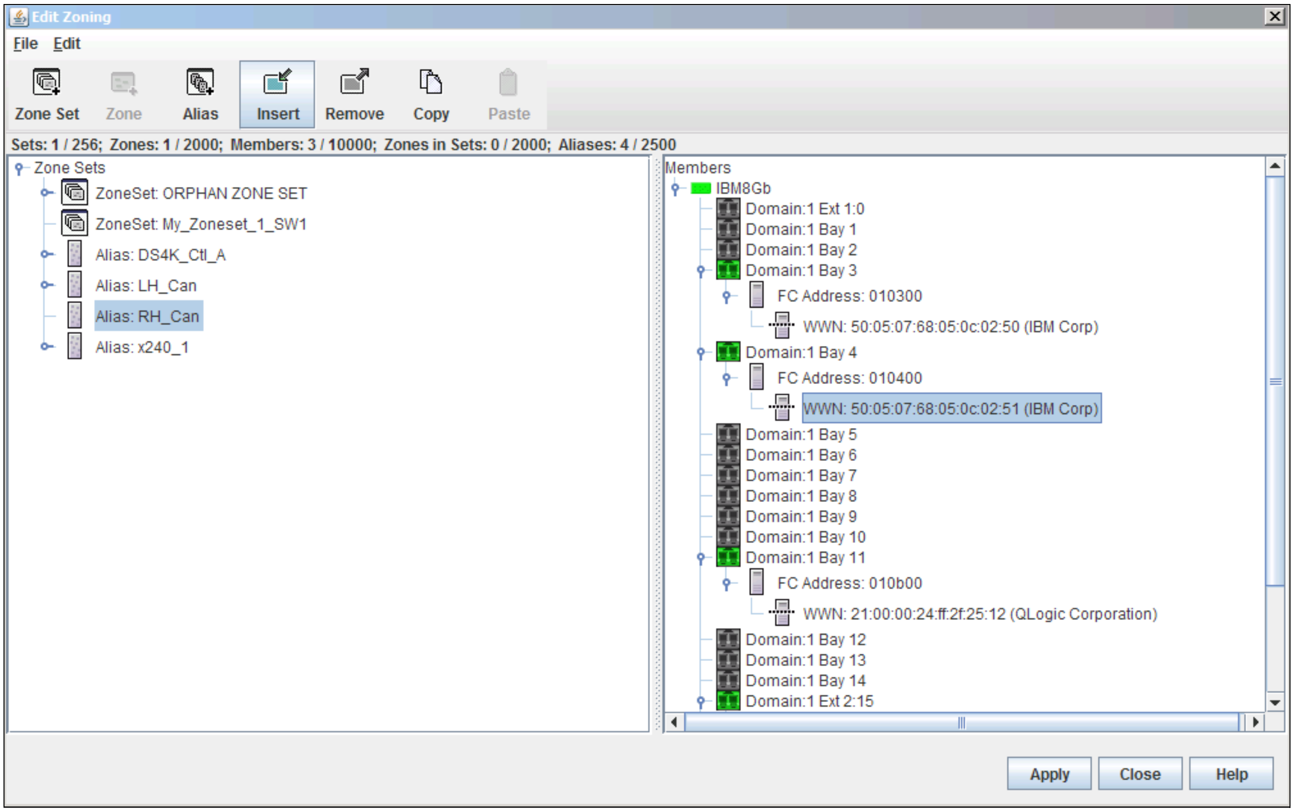


Figure 11-25 Add WWN to alias

When this is completed, you can check that the WWN has been added to the alias as shown in Figure 11-26

Figure 11-26 Check that WWN has been added to alias

Create Zone

We now want to create a zone and add to it aliases that are already created. Figure 11-27 on page 466 shows the main zoning window, highlight the zoneset you wish the zone to belong to and click the 'Zone' square. A window will popup called 'Create Zone' where you can type in the zone name. When completed click the 'OK' box.

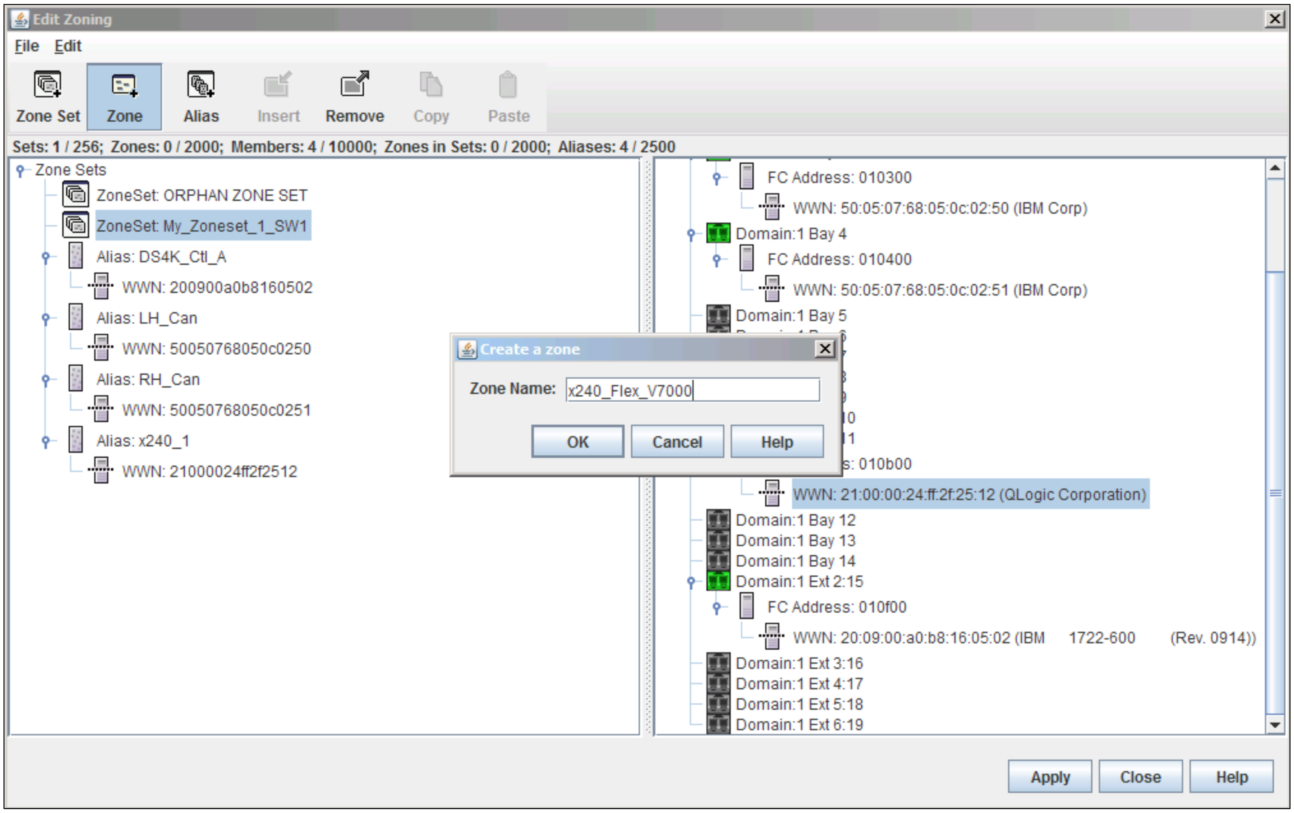


Figure 11-27 Create and name a zone

Adding an alias to a zone

Figure 11-28 shows the highlighted alias where it is dragged using the mouse onto the required zone. This is repeated for all aliases required in that zone.

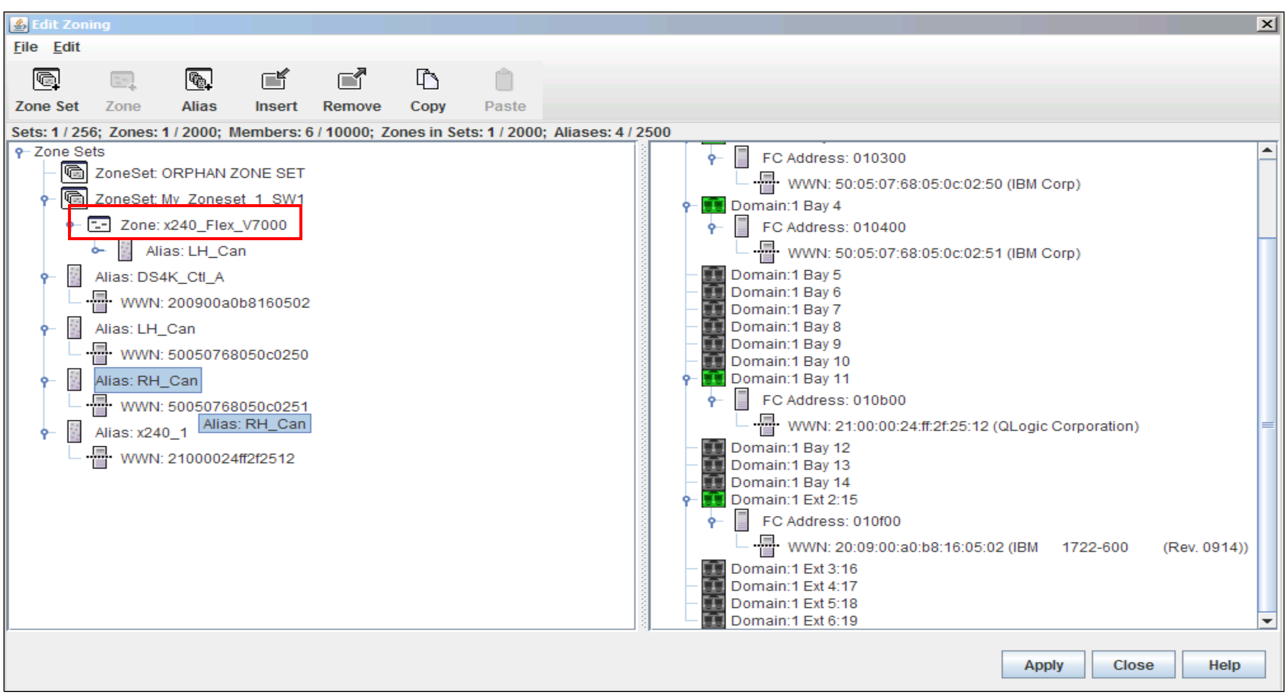


Figure 11-28 Drag alias onto zone name using mouse

Notice that the zone is associated with the required zoneset as specified when the zone was created and named in Figure 11-27 on page 466.

Activate the Zoneset

The aliases and zones have been created and placed in the zoneset, all that is required now is to activate the zoneset. To do this highlight the zone you wish to activate with the mouse and click the 'Apply' box in the bottom right hand corner of the window as shown in Figure 11-29 You will be prompted to save the zoning, click the 'Save Zoning' with the mouse. Continue if you are satisfied with the configuration changes you have made, click the 'Yes' box with the mouse.

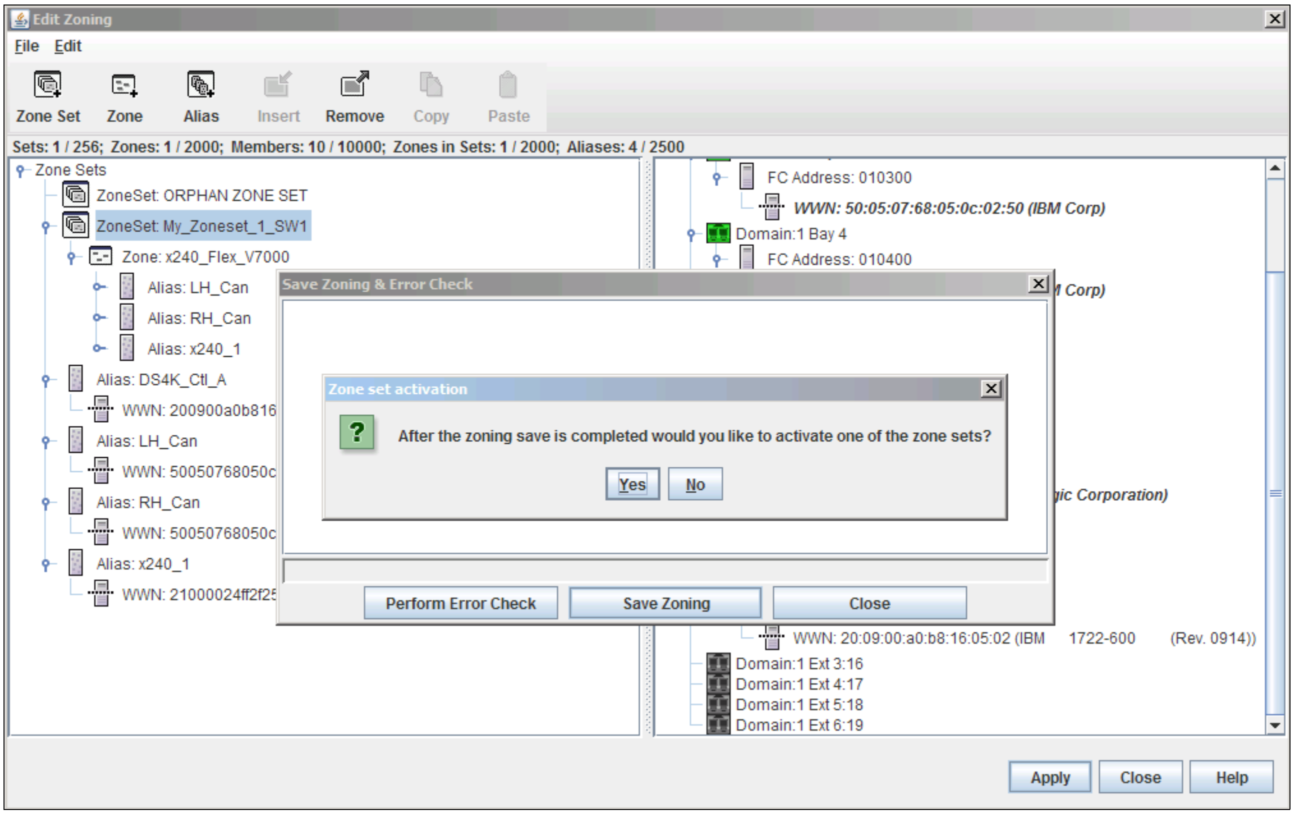


Figure 11-29 Activate zoneset - order of tasks

Figure 11-30 on page 468 will prompt you further for the zoneset name, highlight the required zoneset with the mouse and click on the 'OK' box to continue.

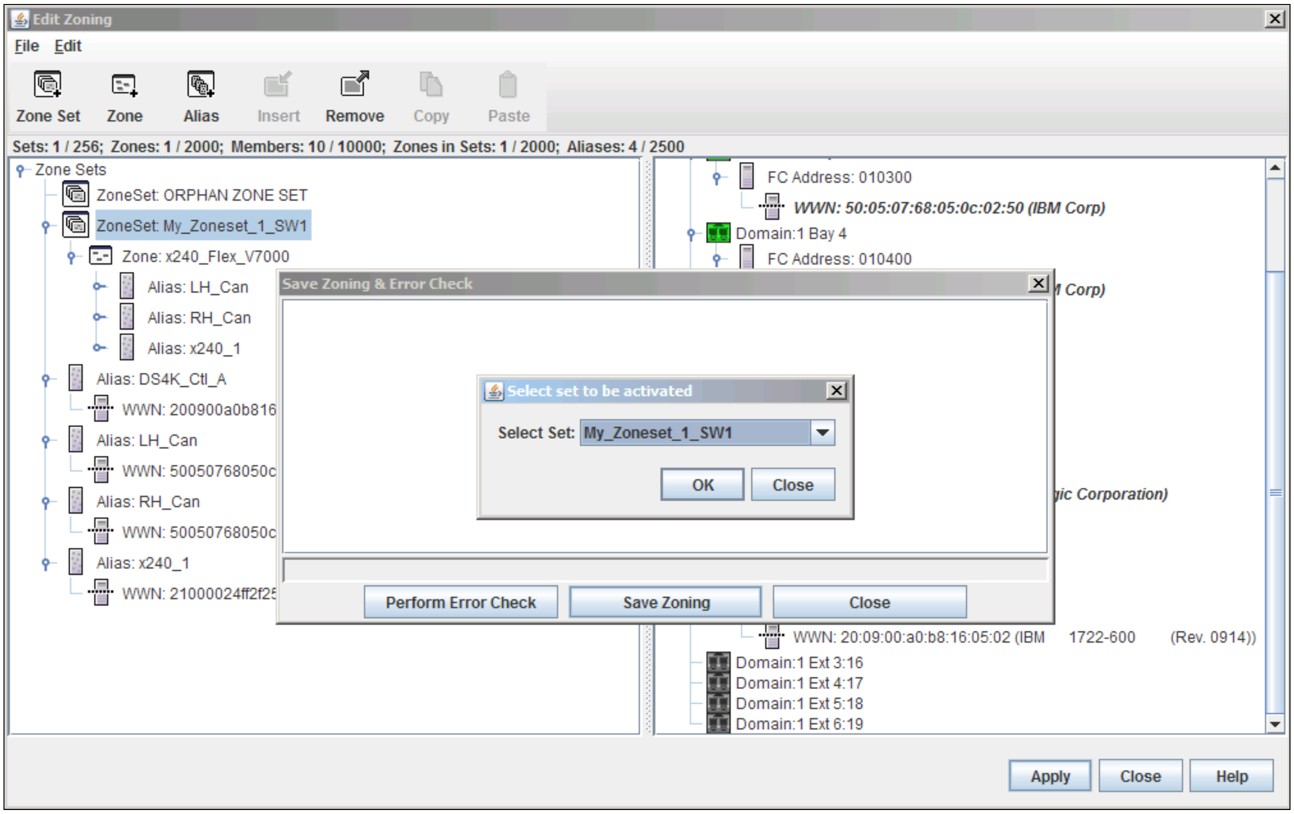


Figure 11-30 Re confirm zoneset activation

When you have completed all zoning, click the ‘Close’ box to the right of the ‘Apply’ box to return to the main switch menu.

This process needs to be repeated for the second FC3171 8Gb SAN Switch in the chassis so as to allow for redundancy by using both fabrics in the chassis.

11.5.4 Multipathing

The above example has created a single zone which includes the initiator from the x240 compute node and two targets, one from each of the storage node canisters. From a compute node host perspective it now has two paths through the FC switch to the IBM Flex System V7000 Storage Node. Repeating the process on the second IBM Flex System Enterprise Chassis FC switch, two more paths are included to the compute node for storage node access. With the compute node host operating system device drivers, it will now have the capability with multipathing to use the four paths available to it. This provides an enhancement for availability and storage performance to the host system on the compute node.

11.5.5 FC Switch Transparent Mode and NPIV

The IBM Flex System FC3171 8Gb SAN Switch used in the zoneset and zone creation example (section 11.5.3, “Zoning a Compute Node for Storage Allocation” on page 460) shows where it has the full functionality of an FC switch. It can connect to an external FC SAN fabric as shown in Figure 11-19 on page 459. The switch can also be configured into ‘Transparent Mode’ where it will connect to an external FC SAN switch using F_Port-N_Port instead of E-Port-E-Port connectivity. From the external SAN and switches all zoning

configuration and management is undertaken. This negates the zoning requirement described previously to be carried out on the FC3171 8Gb SAN Switch. Figure 11-31 shows how to enable ‘Transparent Mode’.

From the main Switch view select ‘Switch’ option in top left of the window then click on ‘Advanced Switch Properties’ with the mouse.

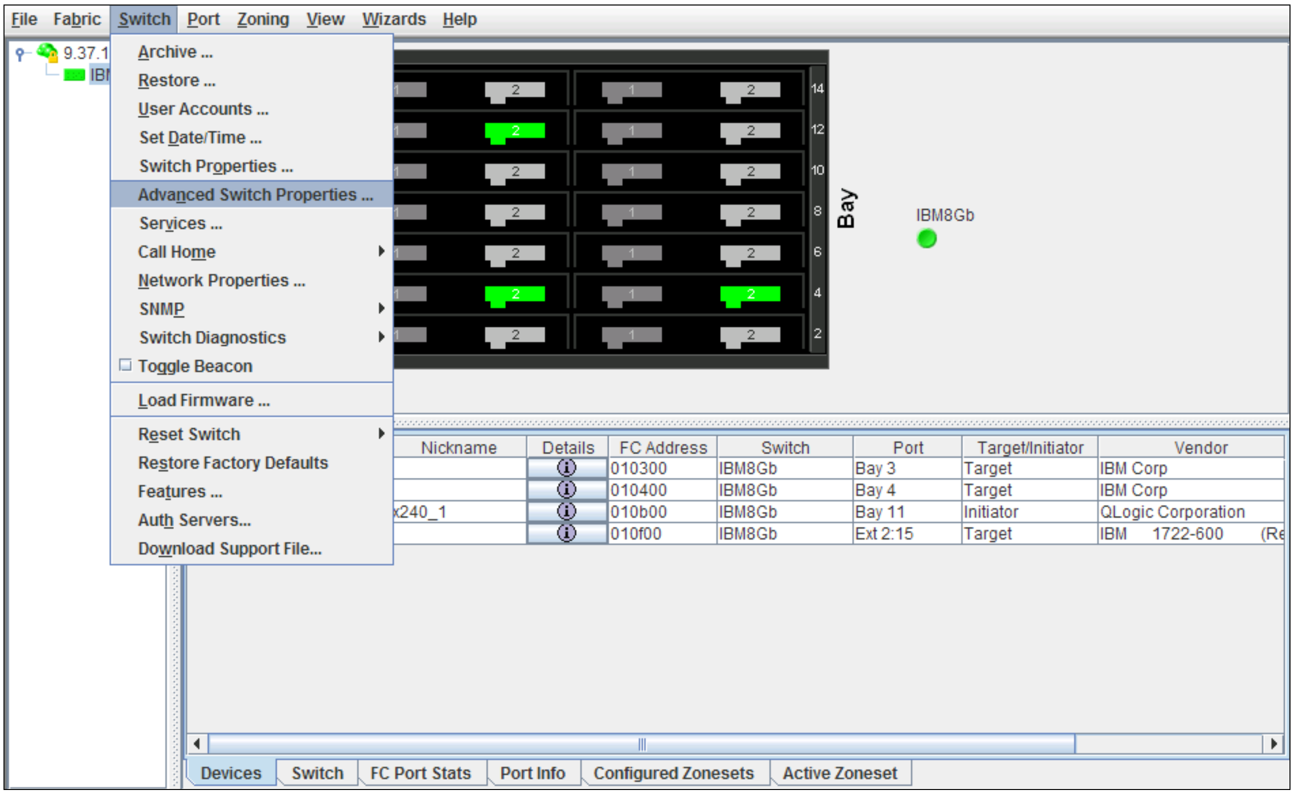


Figure 11-31 Accessing Advanced Switch Properties on FC Switch

Figure 11-32 shows the Advanced Switch Properties window where there is an option to enable Transparent Mode.

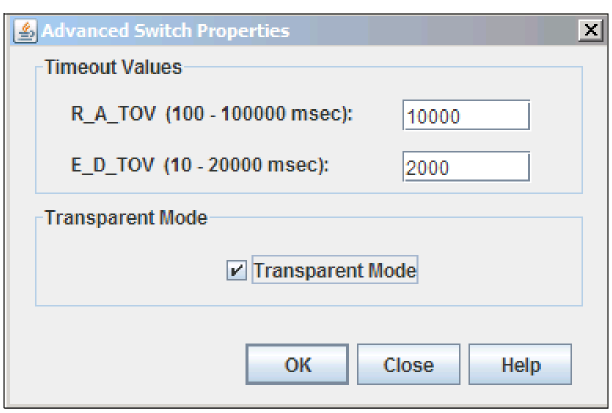


Figure 11-32 Advanced Switch Properties Box

With the mouse, click on the Transparent Mode tick box as shown and click on the ‘OK’ box to complete. As a switch default Transparent Mode is not enabled, if required the above procedure will be required.

N_Port ID Virtualization

In Transparent Mode the FC3171 8Gb SAN Switch becomes a gateway device on FC SAN. A gateway device that uses NPIV technology might be referred to by various names, including NPort concentrator, NPort Aggregator, or NPIV Gateway. Vendors also use many terms to refer to such gateways. For example, Cisco uses the term N_Port Virtualization (NPV), Brocade uses the term Access Gateway, and QLogic uses the term Transparent Mode.

The gateway device employs N_Port ID Virtualization (NPIV) technology which is used in traditional Fibre Channel networks to reduce the number of domains in the overall fabric. The number of FC switches (domains) that can join a single fabric has a theoretical limit of around 250. It also has practical limit generally in the range 50-75. The recommended limits vary by vendor.

In Transparent Mode on the edge of an FC SAN fabric NPIV is used to allow the FC3171 8Gb SAN Switch to appear as a switch (F_Port) to the storage nodes. It will also appear as a node (N_Port) to the external FC switch to which it connects, acting as "concentrator" for multiple N_Port storage nodes. The FC3171 8Gb SAN Switch will allow the attachment of multiple N_Ports of storage nodes but does not use a domain in the external FC fabric.

Using Transparent Mode and NPIV any interoperability issues of merging external SAN switches to the IBM Flex System Enterprise Chassis switch module which may be from different vendors are overcome. Also there is the benefit of easy management as FC3171 8Gb SAN Switch becomes a node in the external FC SAN fabric.

The Transparent Mode option will only be considered when connecting the FC3171 8Gb SAN Switch to an external SAN fabric, as previously mentioned this will be used for the IBM Flex System V7000 Storage Node to connect and manage external storage controllers. Another option is to use the IBM Flex System FC3171 8Gb SAN Pass-thru switch which does not have full switch functionality and operates as a gateway device only using NPIV. The consequences of using Transparent Mode or a gateway device (IBM Flex System FC3171 8Gb SAN Pass-thru) is that all FC traffic between compute nodes and IBM Flex System V7000 Storage Node will be passed externally from the IBM Flex System Enterprise Chassis to a FC switch on the external SAN fabric via the gateway device (or FCF for FCoE) and back into the IBM Flex System Enterprise Chassis.

11.6 Storage Area Network Summary

This chapter has reviewed each type of SAN connectivity between the IBM Flex System V7000 Storage Node and compute nodes, detailing iSCSI over TCP/IP network, FCoE with 10Gb Ethernet, and FC. We have shown that each has capability to connect to an external SAN of same type. The FCoE having the advantage of connecting to external FCoE or FC SAN in FCF or NPV modes with the IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch.

The flexibility and use of Converged Network Adapter (CNA) on a compute node makes FCoE and iSCSI attractive options with its use of 10Gb Ethernet over a single network. However, a number of sources recommend that FCoE should be separated from other Ethernet traffic. This defeats the purpose of combining FC host bus adapter (HBA) and Ethernet network interface card (NIC) in a CNA card. FCoE is a relatively new storage protocol and is increasing in popularity with Ethernet speeds to exceed that of FC, FCoE may well displace FC as the leading block based storage protocol. Use of FCoE within the IBM Flex System Enterprise Chassis where only one FCF in the internal dedicated network is employed, we would expect to see it perform at its optimum (as with iSCSI and FC). In the

Flex System environment storage performance will be critical so expectations of the selected storage protocol will be very high.

Within the IBM Flex System Enterprise Chassis either option would be suitable between storage and compute nodes. Each of the iSCSI, FCoE or FC have their attributes and advantages over another. Performance for example, would be one attribute where separation between the three could be measured, but such a discussion is beyond the scope of this publication. Major determining factors in this choice of SAN will include the external connectivity whether FC or type of Ethernet. We would favour FCoE and FC over iSCSI in the enclosed IBM Flex System Enterprise Chassis environment as a TCP/IP within the protocol layer is unnecessary in this environment.

Regardless of which SAN storage protocol chosen, IBM Flex System V7000 Storage Node adds flexibility and ease of management to Pure Flex enhancing a solid foundation of compute, storage, networking, and software resources.

12

Host Configuration

This chapter describes how to configure hosts to access the volumes that are created on a IBM Flex System V7000 Storage Node.

12.1 Host configuration overview

In this section we provide an overview of connectivity between servers (hosts) and storage with the primary focus on how to configure the hosts.

A host system, as defined here, is an open-systems computer that is connected to storage through LAN or SAN switches, using either Fibre Channel (FC), iSCSI or FCoE protocols. In a IBM Flex System a host is also called a compute node.

IBM Flex SystemCompute nodes are available with either Intel Xeon or IBM Power7 microprocessors. These Compute nodes use integrated network ports or optional network adapters to connect to external devices through the switches or pass-thru modules that are installed in the chassis.

The following types of Compute nodes are available for IBM Flex Systems:

- IBM System x: These Compute nodes come with Intel Xeon microprocessors and provide the function, reliability, and performance of the X-Architecture® systems in a small-form-factor design. They support a variety of Microsoft Windows, Linux, and VMware operating systems.
- IBM Power Systems: These Compute nodes come with IBM Power7 microprocessors and provide the function, reliability, and performance of the Power7 systems in a small-form-factor design. They support a variety of AIX, Linux, and IBM System i® operating systems.

For more information about nodes, refer to IBM Flex System InfoCenter:
http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.common.nav.doc/compute_blades.html

This chapter covers the following operating systems and protocols:

- Windows on IBM X-Architecture Compute nodes
  - Fibre Channel (FC)
  - iSCSI
  - FCoE
- VMware on IBM X-Architecture Compute nodes
  - Fibre Channel
  - iSCSI
- AIX on IBM Power Systems Compute nodes
  - Fibre Channel (FC)
- LINUX on IBM X-Architecture Compute nodes
  - a short introduction

This chapter focuses on Windows, VMware, AIX and Linux. If you attach any other hosts, for example, IBM System i, then you can find the required information in the IBM Flex System V7000 Storage Node Information Center at the following address:

<< Correct the link below>>

http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp

12.2 Discovering the volumes from the host and Multipath settings

This section shows how to discover host mapped volumes from the host

We assume that you have completed all previous steps so that the hosts and the IBM Flex System V7000 Storage Node are prepared:

- Prepare your operating systems for attachment by installing the appropriate multipath driver if needed. Use the System Storage Interoperability Center (SSIC) to check which code levels are supported to attach your host to your storage. SSIC is a web tool that checks the interoperation of host, storage, switches, and multipathing drivers. It can be found at the following address:
`http://ibm.com/systems/support/storage/ssic/interoperability.wss`
- Perform basic volume configuration.
- Create hosts using the GUI.
- Perform host mapping.

Refer to Chapter 6, “Basic Volume and Host Configuration” on page 241 for more information.

In the IBM Flex System V7000 Storage Node GUI, click **Hosts** → **All Hosts** (Figure 12-1).

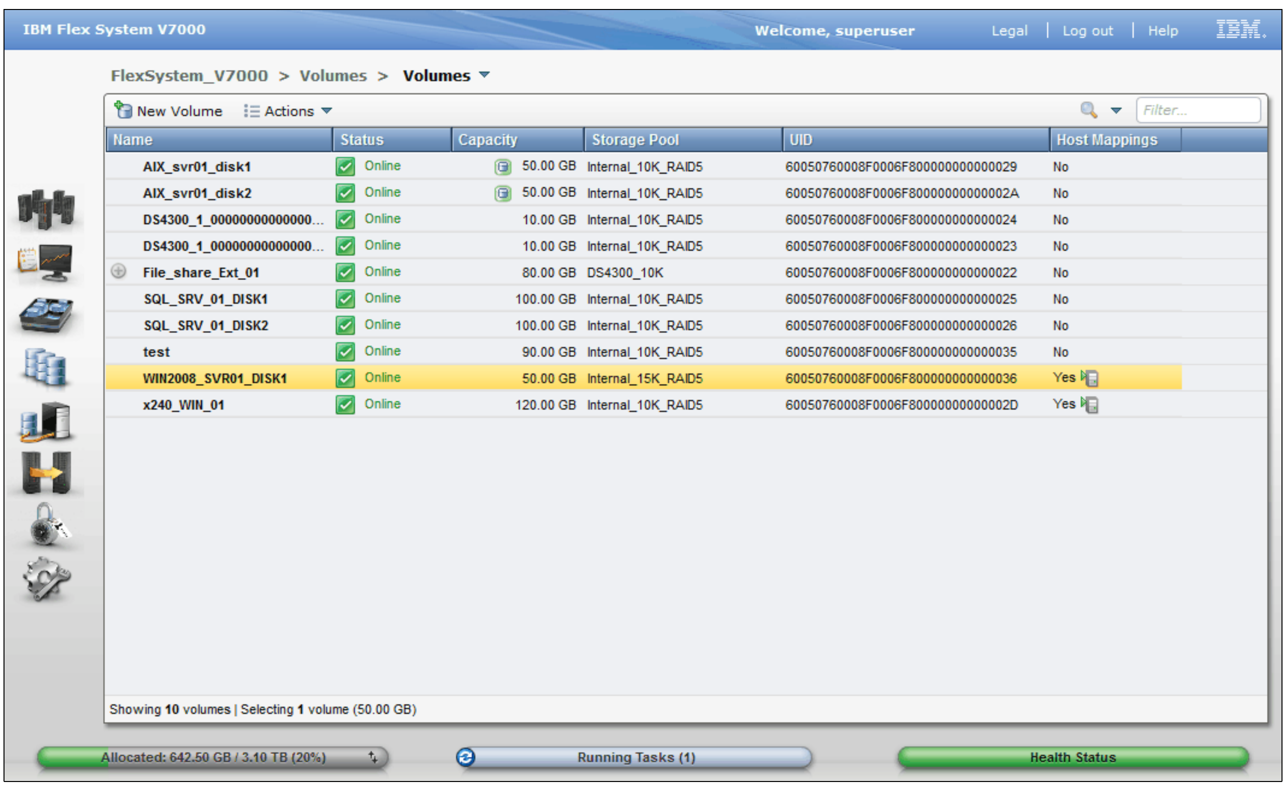


Figure 12-1 Open all hosts

The view that opens gives you an overview of the currently configured and mapped hosts (Figure 12-2).

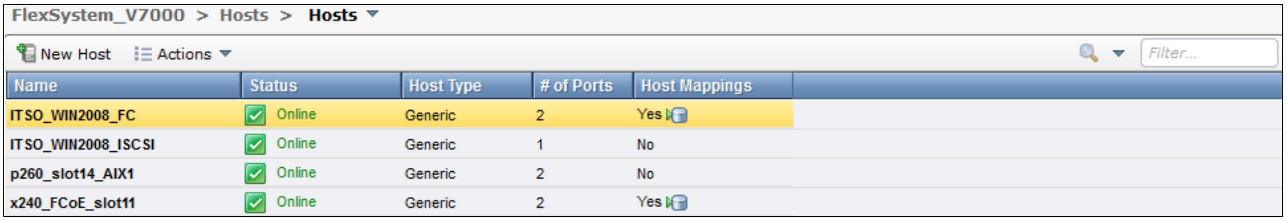


Figure 12-2 All Hosts view

12.3 Windows host attachment

In this section we cover Windows host attachment for Fibre channel (FC), iSCSI and FCoE.

12.3.1 Windows 2008 R2 Fibre Channel volume attachment

The equipment we have for demonstrating volume attachment for an Fibre channel (FC) capable server is an IBM Flex System x240 Compute Node running Microsoft Windows 2008 R2 Enterprise, two IBM Flex System FC3171 8Gb SAN Switch and one IBM Flex System V7000 Storage Node. All is installed in a single IBM Flex System Enterprise Chassis

To perform Fibre Channel volume attachment in Windows 2008 R2, complete the following steps:

1. Right-click your Windows 2008 Fibre Channel host in the All Hosts view (Figure 12-3) and select **Properties**.

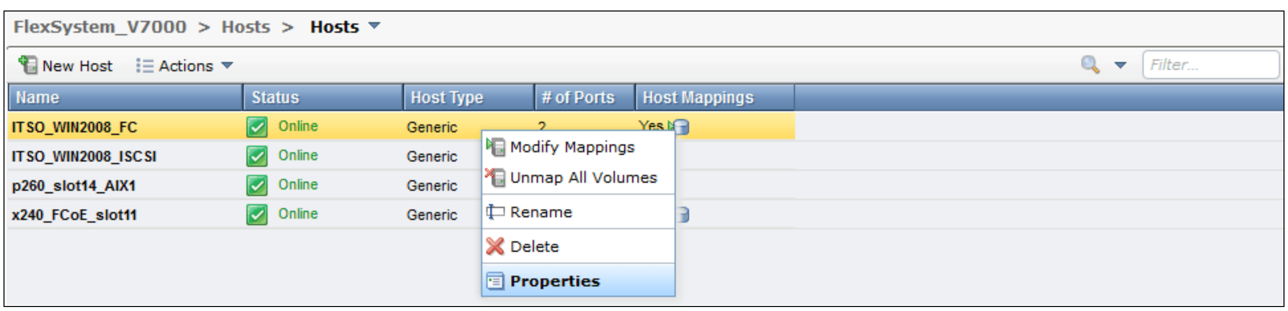


Figure 12-3 Host properties

Navigate to the **Mapped Volumes** tab (Figure 12-4).

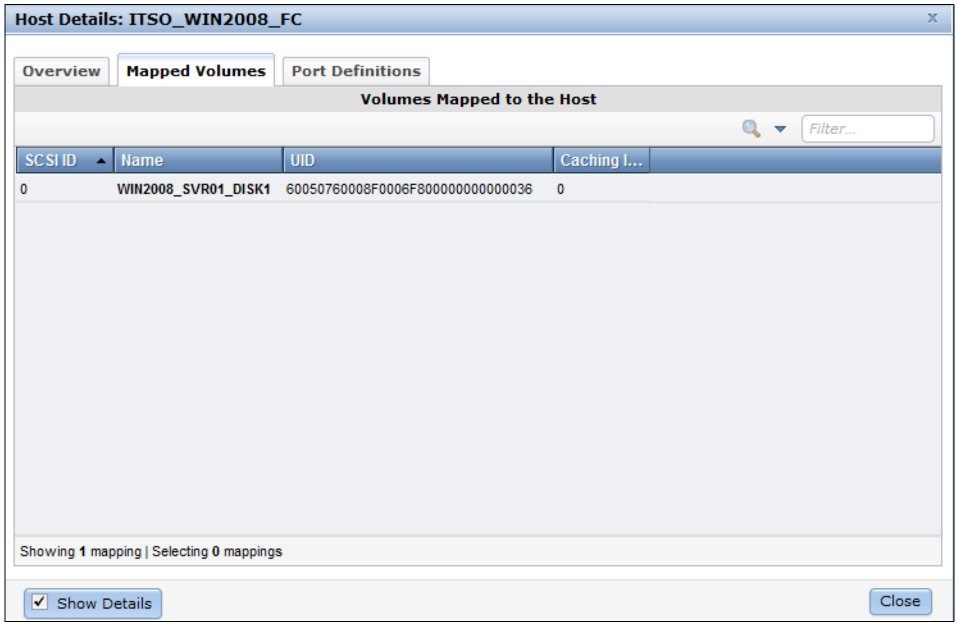


Figure 12-4 Mapped Volumes to a host

The host details show you which volumes are currently mapped to the host, and you also see the volume UID and the SCSI ID. In our example, one volume with SCSI ID 0 is mapped to the host.

2. Log on to your Microsoft host and click **Start** → **All Programs** → **Subsystem Device Driver DSM** → **Subsystem Device Driver DSM**. A command-line interface opens. Enter **`datapath query device`** and press Enter to see if there are IBM Flex System V7000 Storage Node disks connected to this host (Example 12-1).

Example 12-1 datapath query device

```
C:\Program Files\IBM\SDDDSM>datapath query device

Total Devices : 1


DEV#:   0  DEVICE NAME: Disk2 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 60050760008C0006F800000000000028
============================================================================
Path#            Adapter/Hard Disk         State  Mode        Select     Errors
    0      Scsi Port1 Bus0/Disk2 Part0     OPEN   NORMAL           0          0
    1      Scsi Port1 Bus0/Disk2 Part0     OPEN   NORMAL          16          0
    2      Scsi Port2 Bus0/Disk2 Part0     OPEN   NORMAL           0          0
    3      Scsi Port2 Bus0/Disk2 Part0     OPEN   NORMAL          16          0

C:\Program Files\IBM\SDDDSM>
```

3. The output provides information about the connected volumes. In our example, there is one disk connected, Disk 14 for the Windows host, and four paths to the disk are available (State = Open). Open the Windows Disk Management window (Figure 12-5) by clicking **Start** → **Run**, type diskmgmt.msc, and click **OK** or simply open it from Server Manager.

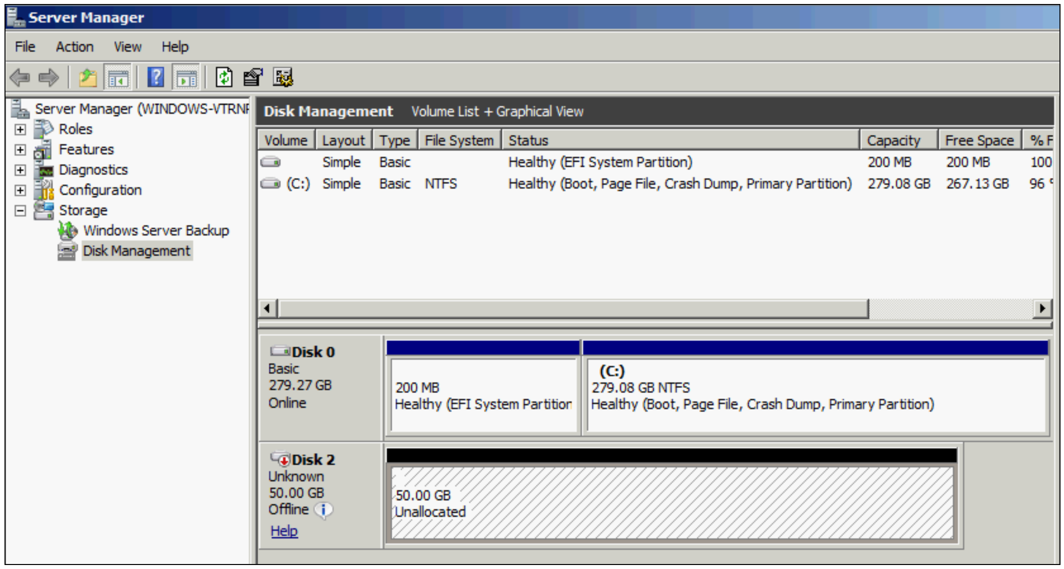


Figure 12-5 Windows Disk Management

Windows device discovery: Usually, Windows discovers new devices, such as disks, by itself. If you have completed all the steps and do not see any disks, click **Actions** → **Rescan Disk** in Disk Management to discover the new volumes.

4. Right-click the disk in the left pane and select **Online** (Figure 12-6).

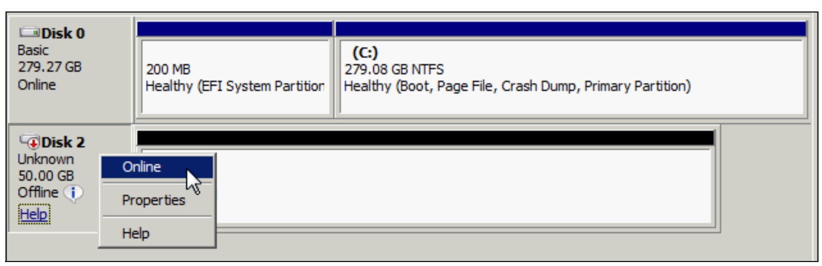


Figure 12-6 Place a disk online

5. Right-click the disk again, click **Initialize Disk** (Figure 12-7), and click **OK**.

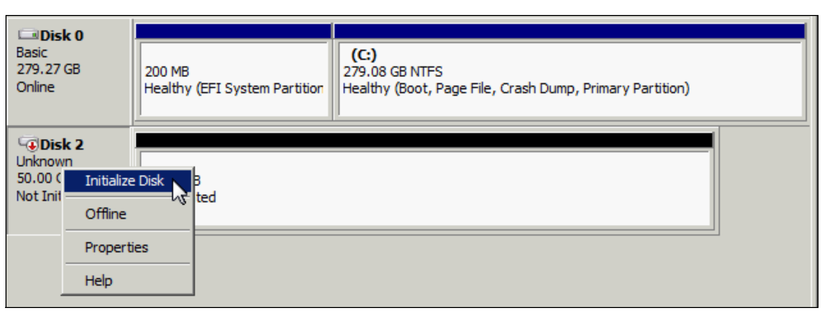


Figure 12-7 Initialize Disk menu

6. Right-click the right pane and click **New Simple Volume** (Figure 12-8).

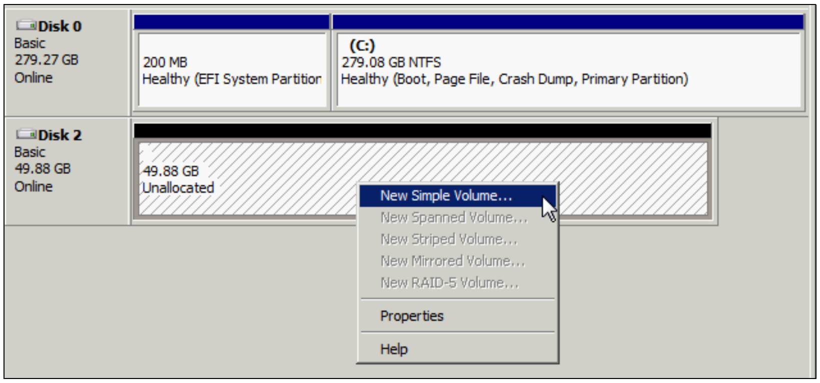


Figure 12-8 New Simple Volume

7. Follow the wizard and the volume is ready to use from your Windows host (Figure 12-9).

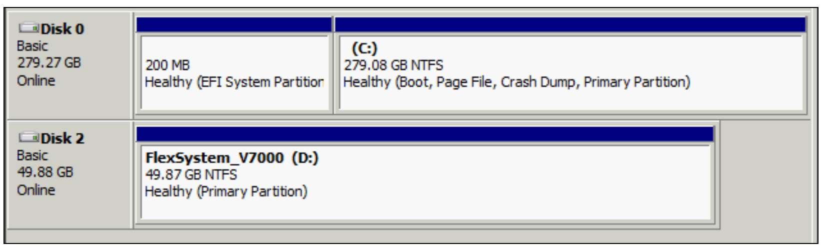


Figure 12-9 Volume is ready to use

The basic setup is now complete, the IBM Flex System V7000 Storage Node is configured, and the host is prepared to access the volumes over several paths and is able to store data on the storage subsystem.

12.3.2 Windows 2008 R2 iSCSI volume attachment

The equipment we have for demonstrating volume attachment for an iSCSI capable server is an IBM Flex System x240 Compute Node running Microsoft Windows 2008 R2 Enterprise, two IBM Flex System Fabric CN4093 10Gb Converged Scalable Switches and one IBM Flex System V7000 Storage Node. All is installed in a single IBM Flex System Enterprise Chassis.

To perform iSCSI volume attachment in Windows 2008, complete the following steps:

1. Right-click your Windows 2008 iSCSI host in the All Hosts view (Figure 12-10), click **Properties**, and navigate to the **Mapped Volumes** tab (Figure 12-11 on page 481).

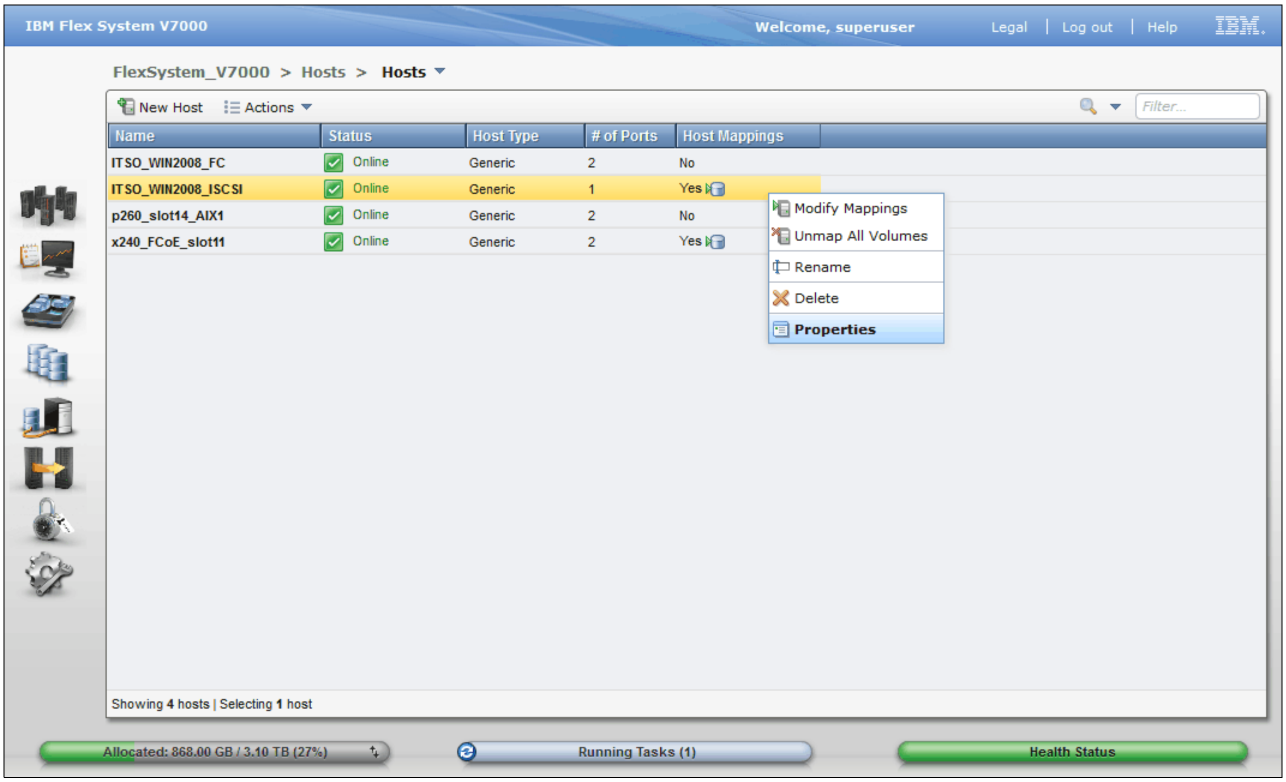


Figure 12-10 All Hosts view - ISCSI host selected

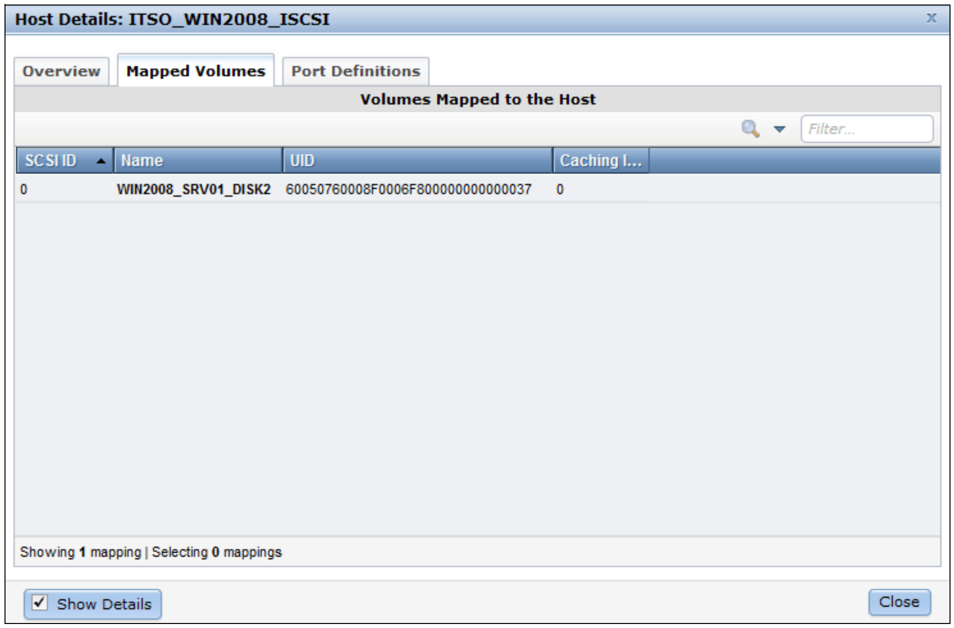


Figure 12-11 Mapped Volumes to an iSCSI host

The host details show you which volumes are currently mapped to the host, and you also see the volume UID and the SCSI ID. In our example, one volume with SCSI ID 0 is mapped to the host.

2. Log on to your Windows 2008 host and click **Start** → **Administrative Tools** → **iSCSI Initiator** to open the iSCSI Configuration tab (Figure 12-12 on page 482).

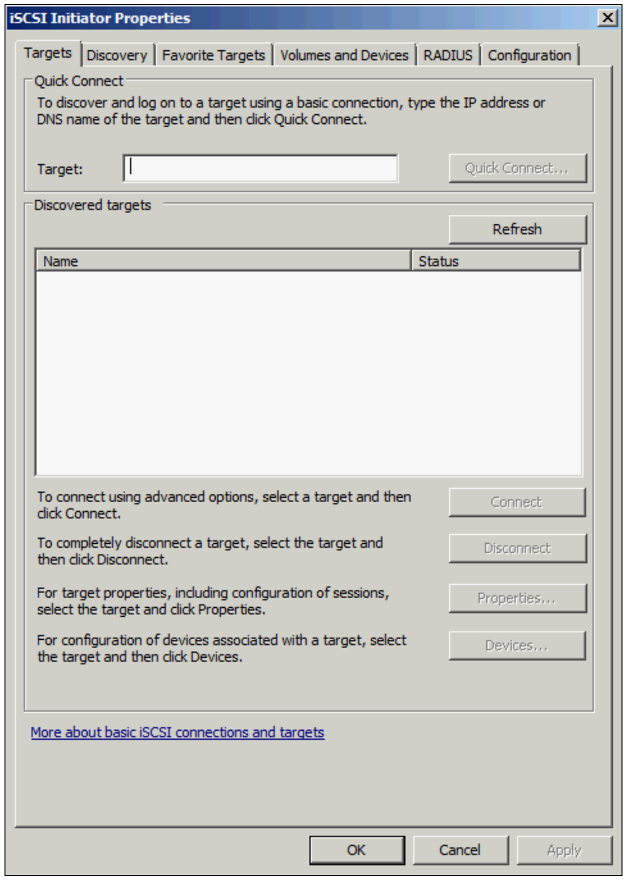


Figure 12-12 Windows iSCSI Configuration tab

3. Enter the IP address of one of the IBM Flex System V7000 Storage Node iSCSI ports and click **Quick Connect** (Figure 12-13).

iSCSI IP addresses: The iSCSI IP addresses are different for the cluster and canister IP addresses. The iSCSI addresses are configured directly on each of the canisters. In case of a failover caused by a failing canister the iSCSI IP-addresses fail over to the surviving canister and therefore always remain online.

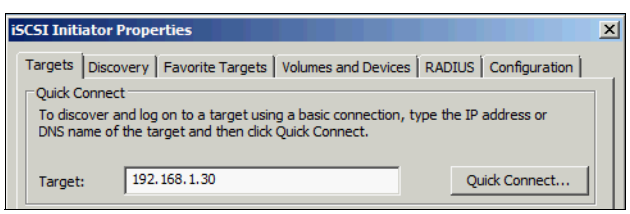


Figure 12-13 iSCSI Quick Connect

The IBM Flex System V7000 Storage Node initiator is discovered and connected. (Figure 12-14).

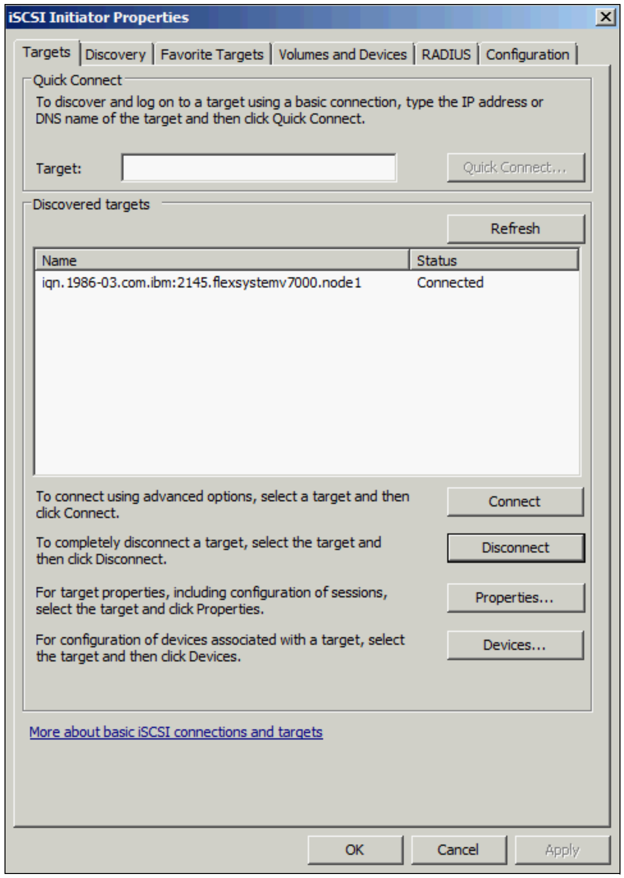


Figure 12-14 iSCSI Initiator target is connected

Now you have completed the steps to connect the storage disk to your iSCSI host, but you are only using a single path at the moment. To enable multipathing for iSCSI targets, more actions are required. Complete the following steps:

1. Click **Start** → **Run** and type cmd to open a command prompt. Enter **`ServerManagerCMD.exe -install Multipath-IO`** and press Enter (Example 12-2).

Example 12-2 Installing MPIO

```
C:\Users\Administrator>ServerManagerCmd.exe -Install Multipath-IO

Start Installation...
[Installation] Succeeded: [Multipath I/O] Multipath I/O.
<100/100>

Success: Installation succeeded.
```

2. Click **Start** → **Administrative Tools** → **MPIO**, click the **Discover Multi-Paths** tab, and select the **Add support for iSCSI devices** check box (Figure 12-15).

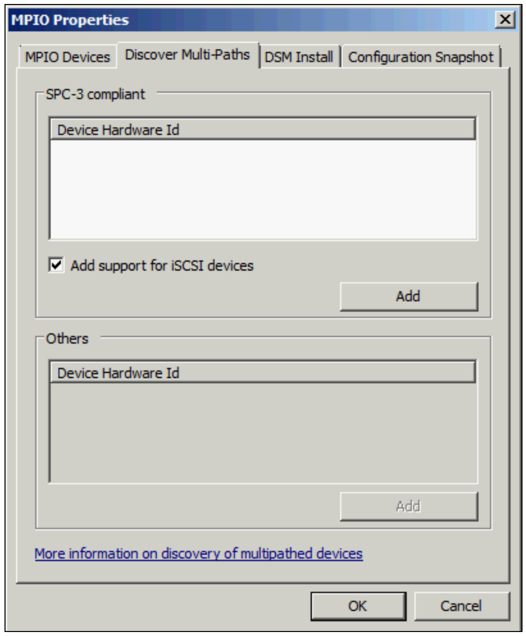


Figure 12-15 Enable iSCSI MPIO

3. Click **Add** and confirm the prompt to reboot your host.
4. After the reboot, log on again, click **Start** → **Administrative Tools** → **iSCSI Initiator** to open the iSCSI Configuration tab, and navigate to the **Discovery** tab (Figure 12-16 on page 485).

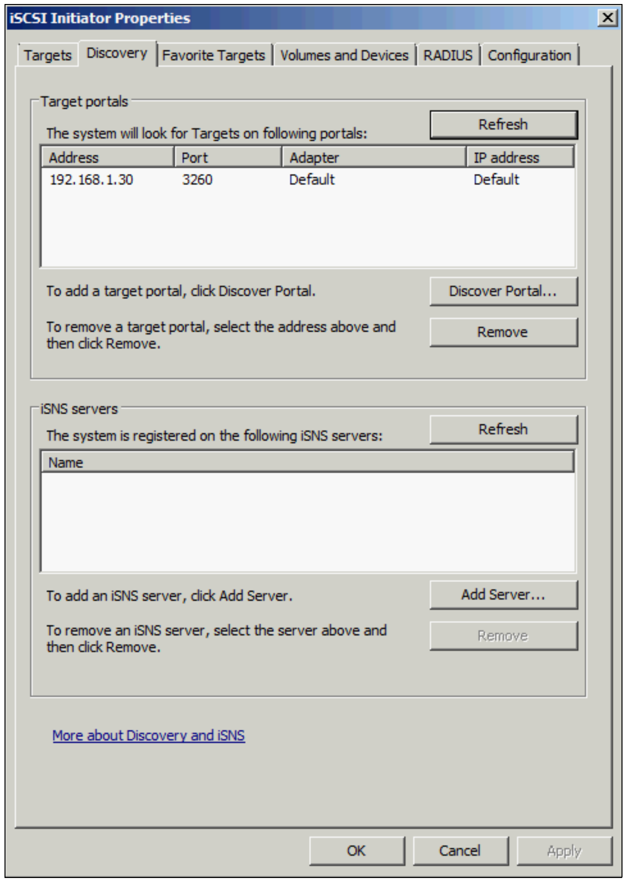


Figure 12-16 iSCSI Properties Discovery tab

5. Click **Discover Portal...**, enter the IP address of another IBM Flex System V7000 Storage Node iSCSI port (Figure 12-17), and click **OK**.

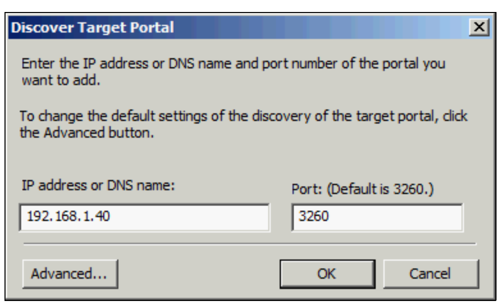


Figure 12-17 Discover Target Portal window

6. Return to the Targets tab (Figure 12-18) and you find the new connection there is listed as Inactive.

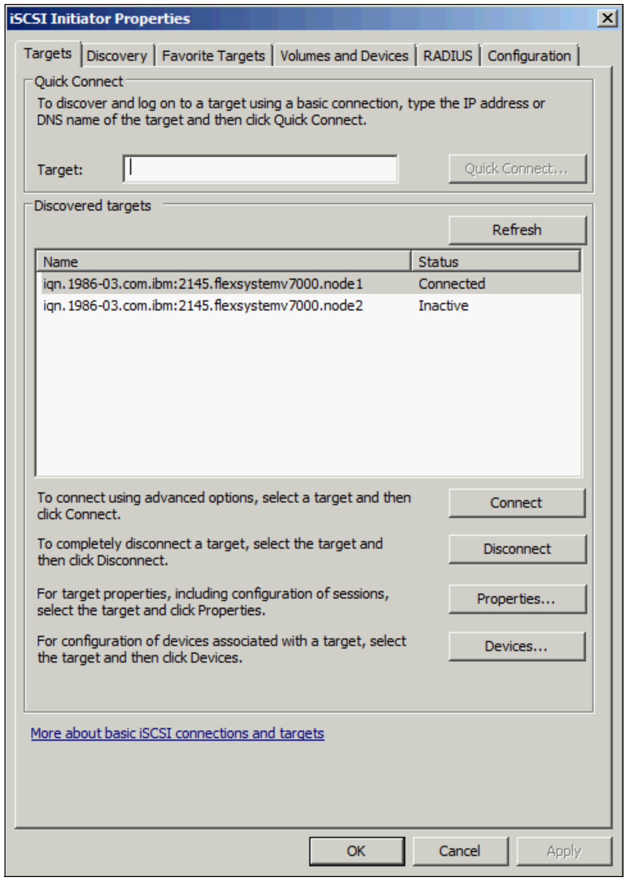


Figure 12-18 Inactive target ports

7. Highlight the inactive port and click **Connect**. The Connect to Target window opens (Figure 12-19).

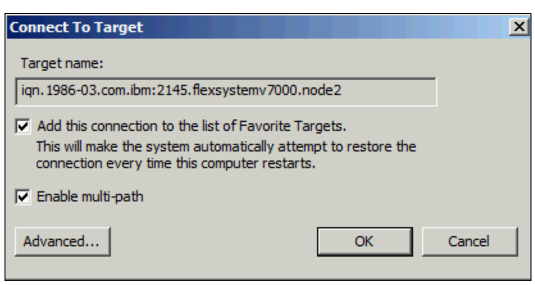


Figure 12-19 Connect to a target

8. Select the **Enable Multipath** check box and click **OK**. The second port now appears as Connected (Figure 12-20).

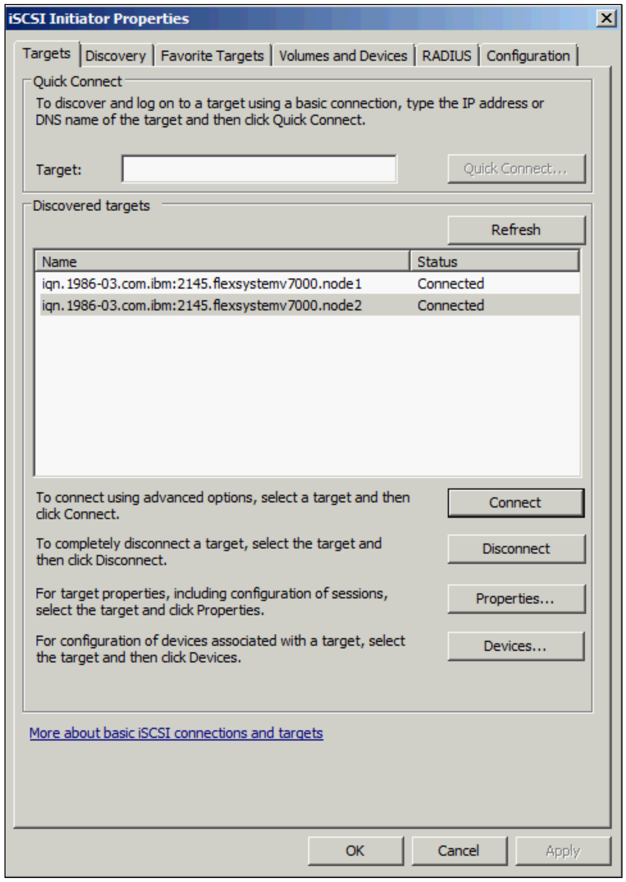


Figure 12-20 Second target port connected

Repeat this step for each IBM Flex System V7000 Storage Node port you want to use for iSCSI traffic. It is possible to have up to four port paths to the system.

9. Open the Windows Disk Management window (Figure 12-21 on page 488) by clicking **Start** → **Run**, typing diskmgmt.msc or simply open it from the Server Manager - click **OK**.

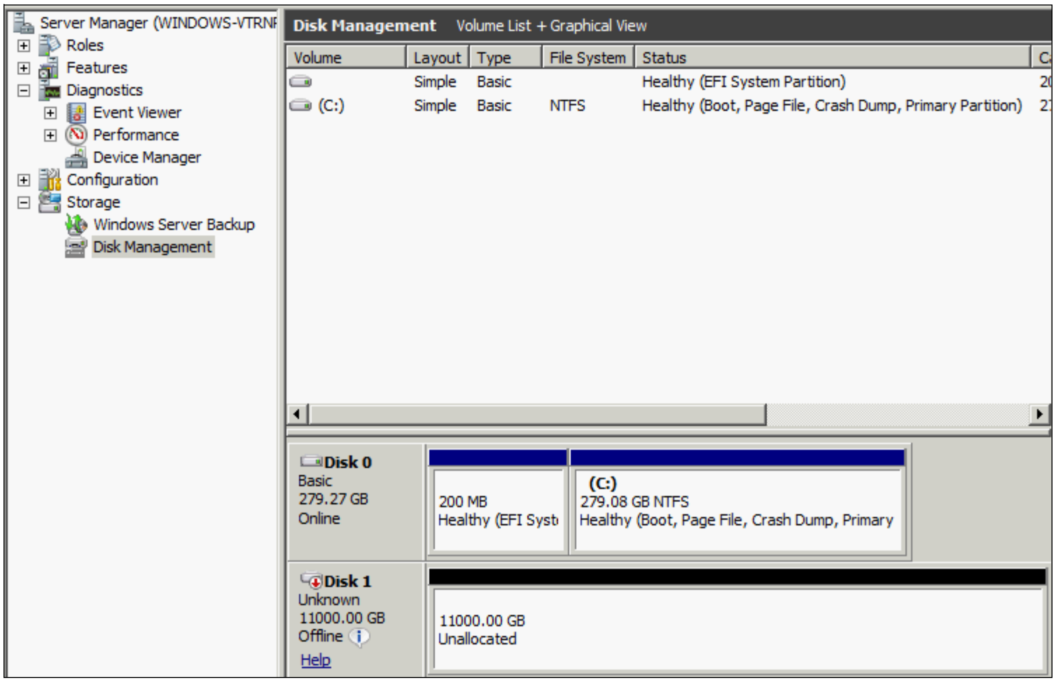


Figure 12-21 Windows Disk Management

10. Set the disk online, initialize it, create a file system on it, and then it is ready to use. The detailed steps of this process are the same as described in 12.3.1, “Windows 2008 R2 Fibre Channel volume attachment” on page 476.

Now the disk is ready for use (Figure 12-22). In our example, we have mapped a 11 TB disk, which is thin-provisioned on the IBM Flex System V7000 Storage Node, to a Windows 2008 Server host using iSCSI.

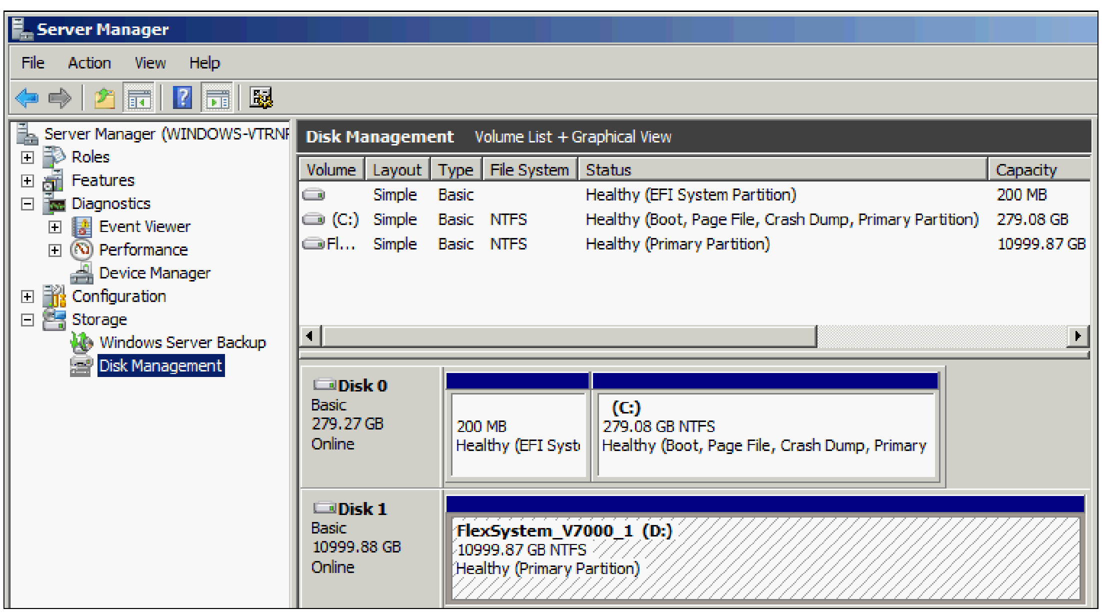


Figure 12-22 Windows Disk Management: Disk is ready to use

The storage disk is available and ready to use for your Windows 2008 R2 server using iSCSI attachment.

Verifying Microsoft iSCSI SW initiator setup

From the Windows server iSCSI Initiator Applet menu Figure 12-23 shows the IBM Flex System V7000 Storage Node canister IQNs defined.

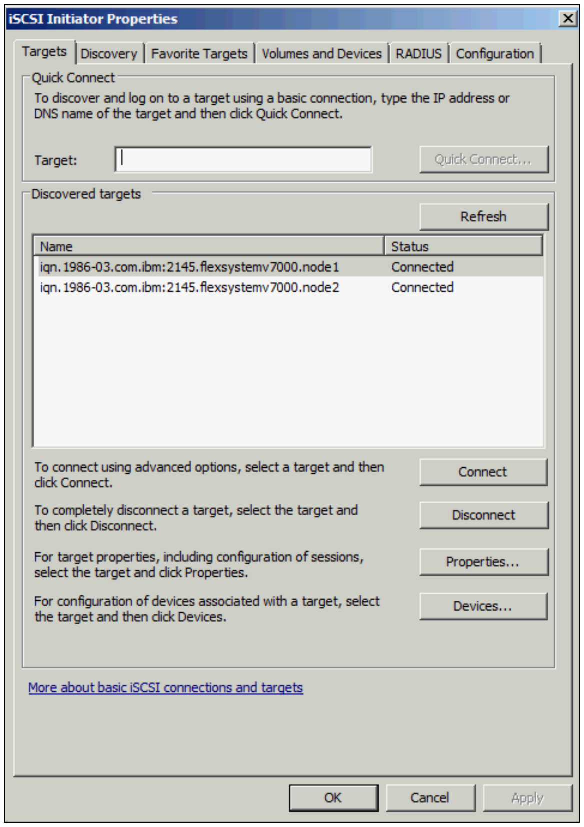


Figure 12-23 Canister IQN define in iSCSI initiator

From this window click on the 'Properties' box with the mouse for node1 highlighted as shown in Figure 12-23. Figure 12-24 will show both canister IP addresses for node1 from the 'Portal Groups' tab.

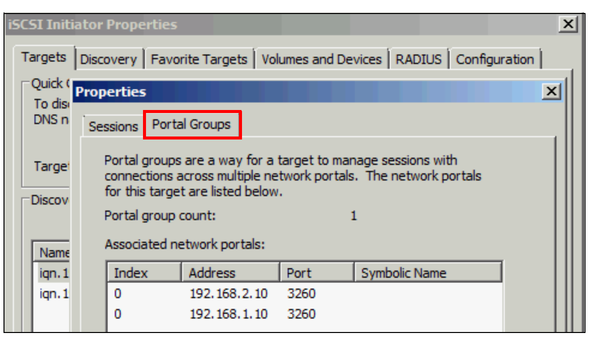


Figure 12-24 Portal Groups tab shows both IP addresses of canister

The 'Discovery' tab in the 'Properties' as shown in Figure 12-25 on page 490 shows all the portal target IP addresses of the iSCSI configuration.

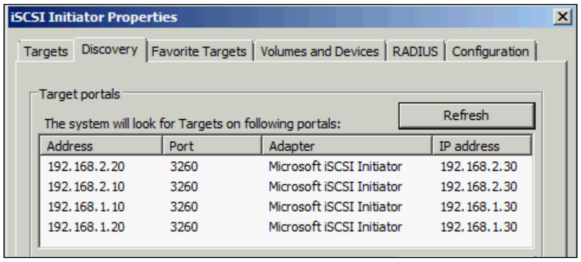


Figure 12-25 iSCSI target IP addresses and adapter portal IPs

The 'Favorite Targets' tab shows two entries for each canister as shown in Figure 12-26 where details of one target is shown, these correspond with the target and portal IP addresses as shown in Figure 12-25.

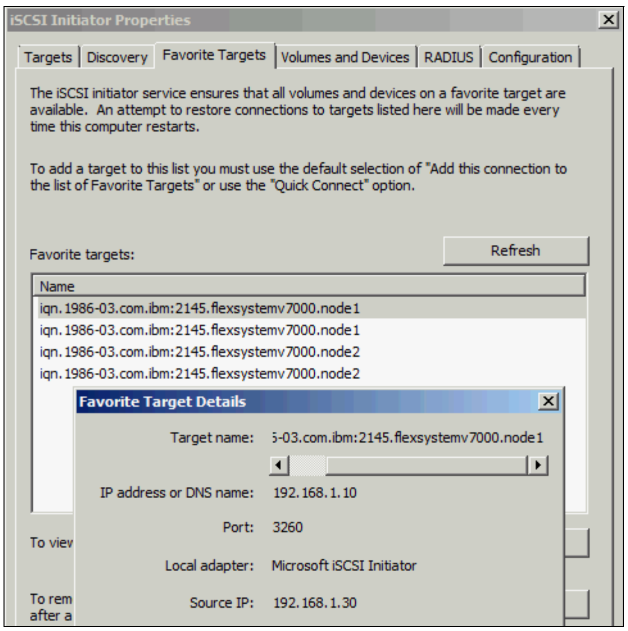


Figure 12-26 Favorite Target details

We then added a single device and then had a look at the properties of that device for each iSCSI target. In Figure 12-27 on page 491, the first target was highlighted and 'Devices' box clicked.

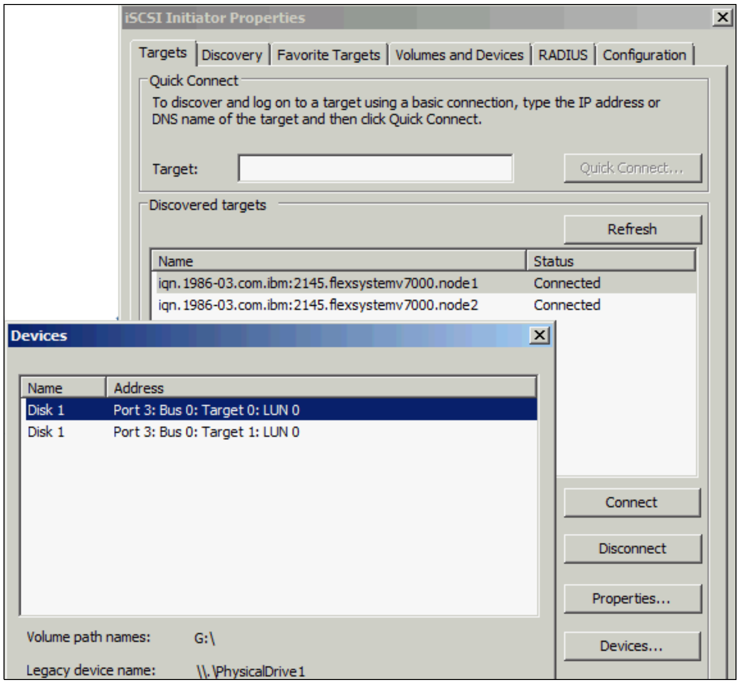


Figure 12-27 Two 'Disk 1's identified on first target

From this there are two disks of the same name with identified as 'Target 0' and 'Target 1', these represent two paths, one from each x240 host adapter to the node canister on separate networks.

From the other iSCSI target of node 2 is shown in Figure 12-28

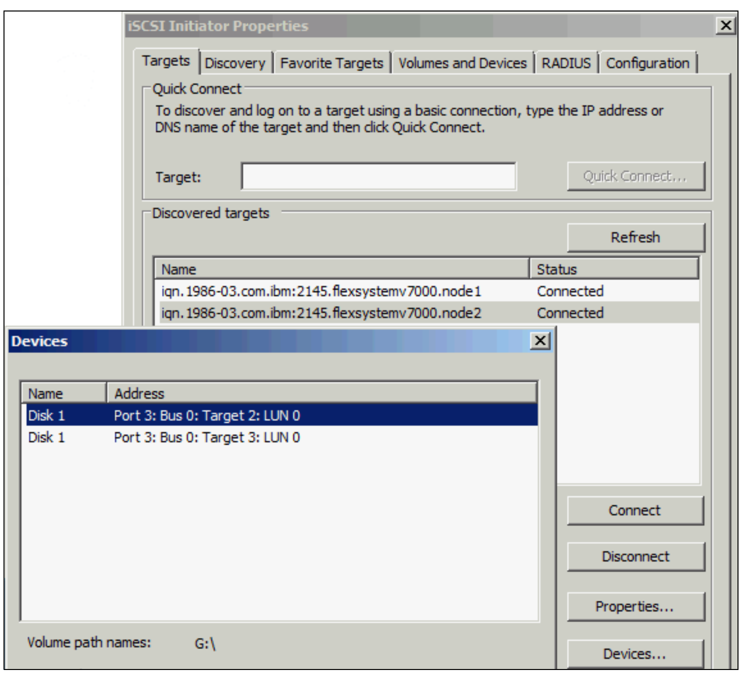


Figure 12-28 Two more devices on second target

On the second iSCSI target, 'Disk 1' has two more devices 'Target 2' and Target 3' confirming four paths to the mapped LUN from the IBM Flex System V7000 Storage Node. Once the

LUN was mapped the ‘Auto Configure’ box was clicked from the ‘Volumes and Devices’ tab as shown in Figure 12-29

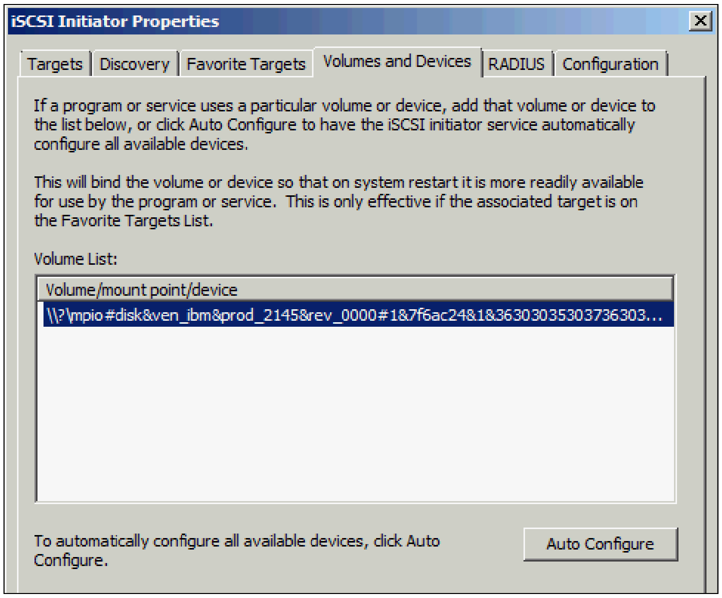


Figure 12-29 Single device identified in tab.

This completes the verification of the iSCSI setup.

iSCSI authentication

Authentication of the host server from the IBM Flex System V7000 Storage Node system is optional and is disabled by default.

The user can choose to enable Challenge Handshake Authentication Protocol (CHAP) authentication, which involves sharing a CHAP secret between the IBM Flex System V7000 Storage Node system and the host.

The IBM Flex System V7000 Storage Node as authenticator sends a challenge message to the specific server (peer). The server responds with a value that is checked by the IBM Flex System V7000 Storage Node. If there is a match, the IBM Flex System V7000 Storage Node acknowledges the authentication. If not, the IBM Flex System V7000 Storage Node will terminate the connection and will not allow any I/O to volumes.

A CHAP secret can be assigned to each IBM Flex System V7000 Storage Node host object. The host must then use CHAP authentication to begin a communications session with a node in the system. A CHAP secret can also be assigned to the system.

Because iSCSI can be used in networks where data security is a concern, the specification allows for separate security methods. You can set up security, for example, through a method such as IPSec, which is transparent for higher levels such as iSCSI because it is implemented at the IP level. Details regarding securing iSCSI can be found in RFC3723, *Securing Block Storage Protocols over IP*, which is available at this website:

http://tools.ietf.org/html/rfc3723

12.3.3 Windows 2008 R2 FCoE volume attachment

The equipment we have for demonstrating volume attachment for an FCoE capable server is an IBM Flex System x240 Compute Node running Microsoft Windows 2008 R2 Enterprise, two IBM Flex System Fabric CN4093 10Gb Converged Scalable Switches and one IBM Flex System V7000 Storage Node. All is installed in a single IBM Flex System Enterprise Chassis.

To perform FCoE volume attachment in Windows 2008 R2 complete the following steps:

1. Configure connectivity between the host and the IBM Flex System V7000 Storage Node. This is done on the FCoE capable switches and includes configuring VLANs. For more information about how to configure VLANs on the CN4093 FCoE switches refer to Chapter 11, "SAN Connections and Configuration" on page 431.
2. Configure appropriate zoning for interconnecting the 10Gb FCoE host ports with the IBM Flex System V7000 Storage Node 10Gb FCoE ports. For more information about how to configure zoning on the CN4093 FCoE switches refer to Chapter 11, "SAN Connections and Configuration" on page 431.
3. On the IBM Flex System V7000 Storage Node click your Windows 2008 R2 FCoE host in the All Hosts view (Figure 5-37) and review that a volume is mapped to the host.

   Figure 12-30 shows that a 200GB volume is mapped to our FCoE host.

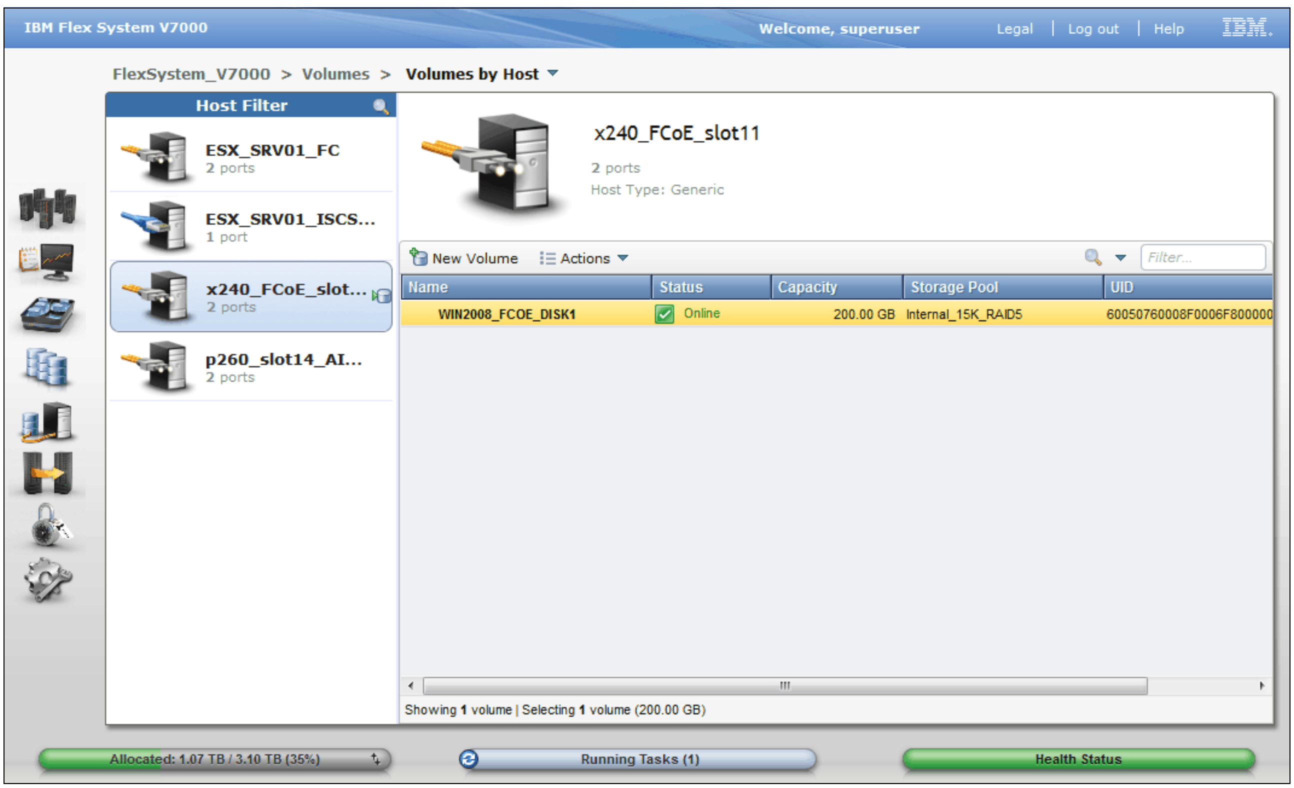


Figure 12-30 Volumes by Host - 200GB volume mapped

You may also navigate to the Hosts screen, right-click your FCoE host and select **Properties** → **Mapped Volumes** for the host. This screen displays which Fibre Channel ports are defined for the host. FCoE ports displays as if they were ordinary FC ports showing World Wide Names (WWNs) for the Host Bus Adapters (HBAs).

Figure 12-31 on page 494 shows the ports that are defined for the FCoE capable IBM Flex System x240 Compute Node.

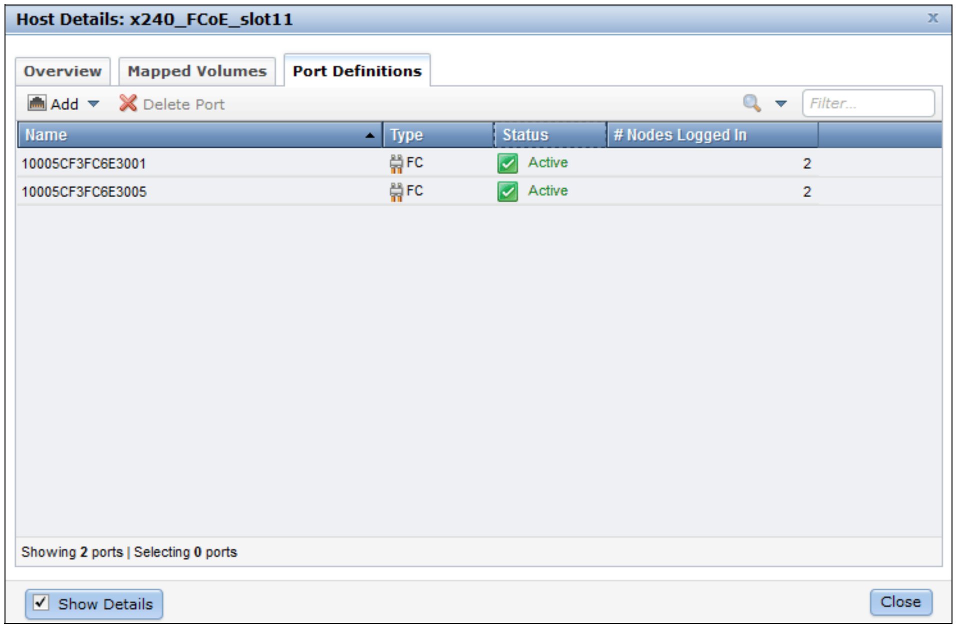


Figure 12-31 FCoE port definitions

4. On the Windows 2008 server make sure appropriate drivers are installed for FCoE attachment. This can be done through the IBM UpdateXpress System Update Installer which can be downloaded from:

   `http://www-947.ibm.com/support/entry/portal/docdisplay?lndocid=SERV-XPRESS`

   Figure 12-32 on page 495 shows how FCoE drivers are installed using the IBM UpdateXpress System Update Installer. IBM UpdateXpress System Update Installer is capable of detecting and recommending which drivers are needed for the Windows 2008 R2 server.

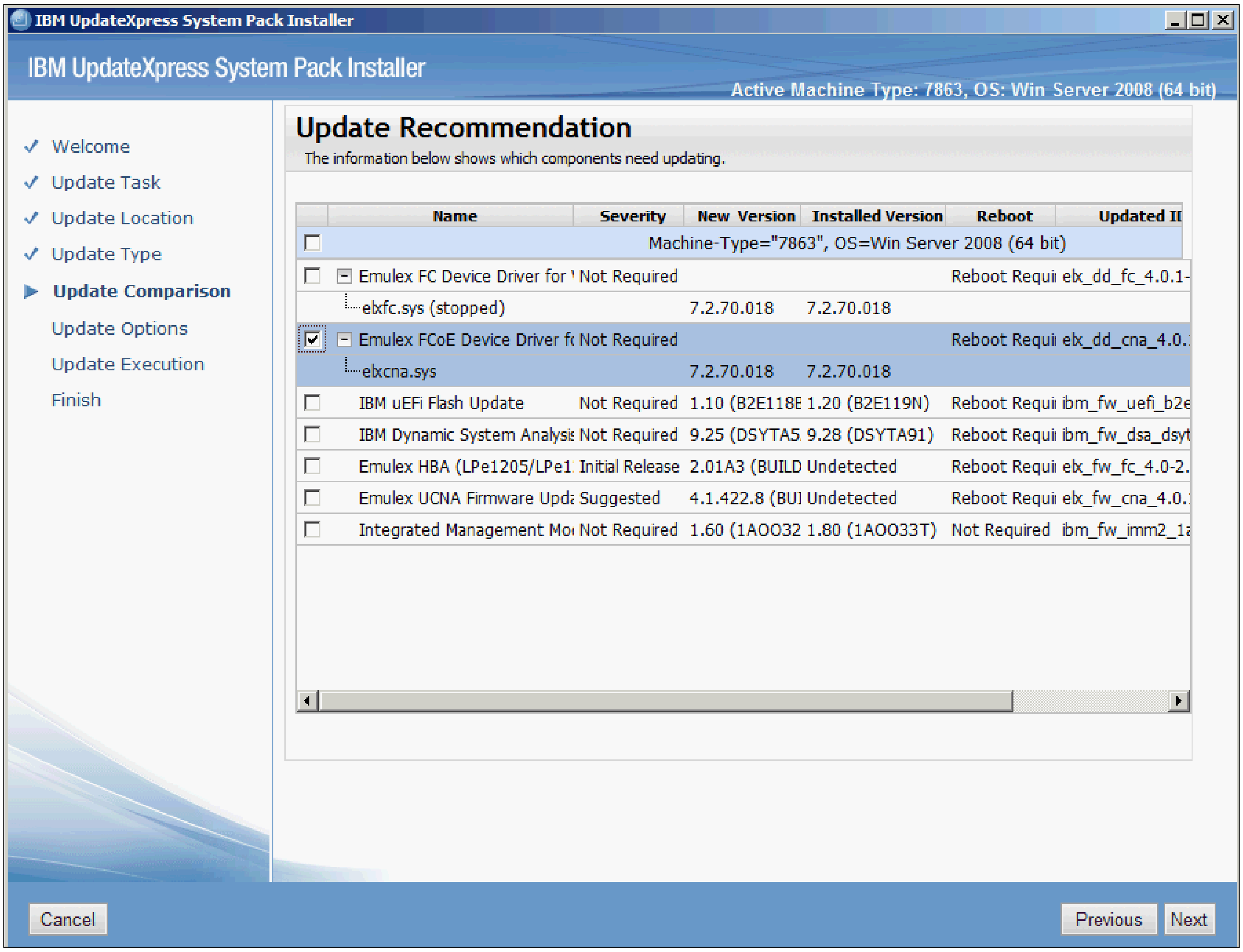


Figure 12-32 FCoE drivers are installed

5. On the Windows 2008 R2 server install appropriate multipathing software. Multipathing software for FCoE is the same as for Fibre Channel (FC) attached hosts and is called Subsystem Device Driver Device Specific Module (SDDDSM). SDDDSM software can be downloaded from:

   `https://www.ibm.com/storage/support`

6. On the Windows 2008 R2 server you may optionally install the Emulex OneCommand Manager which provides detailed information about the configured FCoE ports and their status. Emulex OneCommand Manager can be downloaded from:

   `http://www.emulex.com/downloads/ibm/lightpulse-fc-hba-drivers.html`

   Figure 12-33 on page 496 shows an example of how Emulex OneCommand Manager display the FCoE ports and the volumes it sees when IBM Flex System V7000 Storage Node volumes are mapped to the Windows 2008 R2 server.

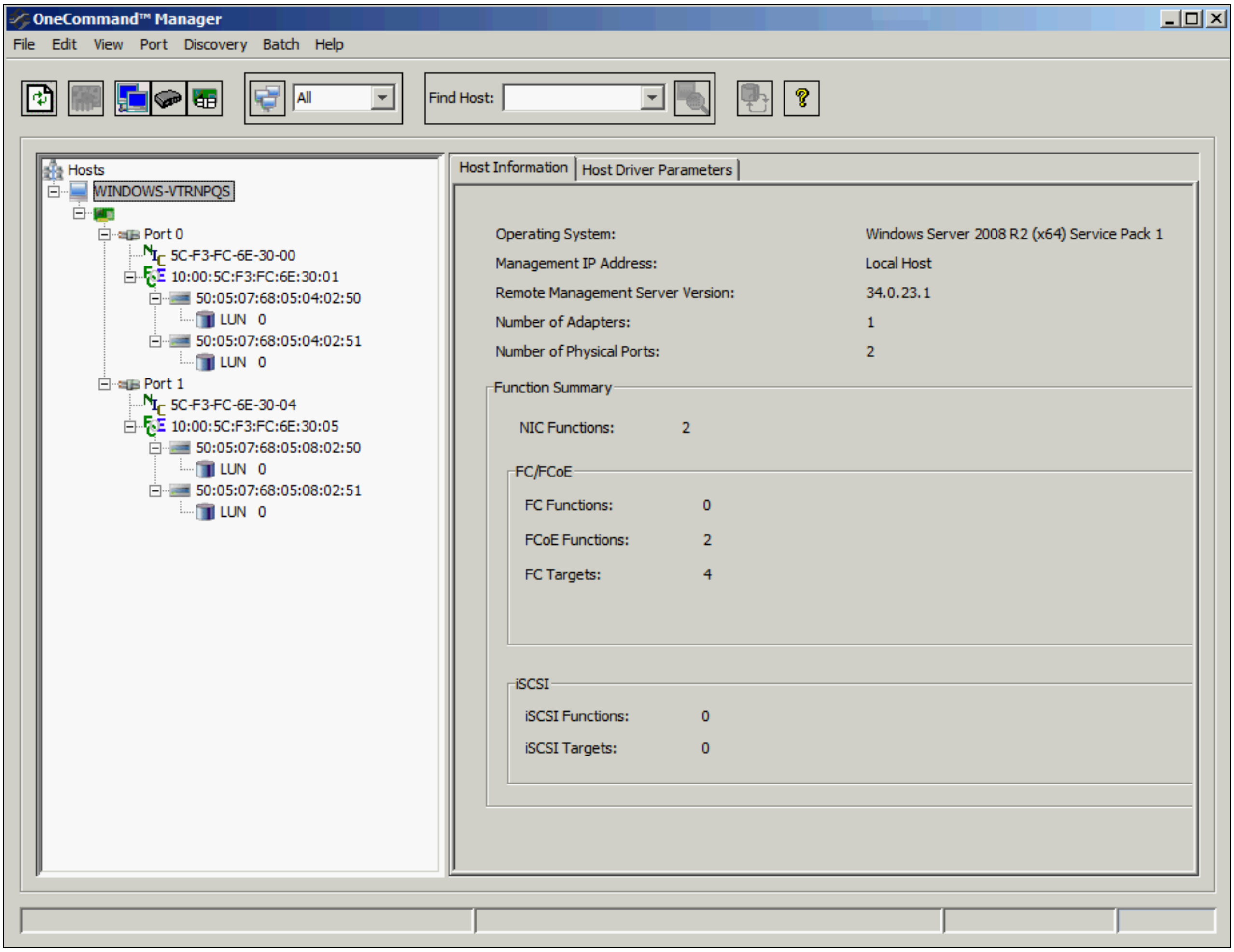


Figure 12-33 Emulex OneCommand Manager

7. On the Windows 2008 R2 server click **Start** → **All Programs** → **Subsystem Device Driver DSM** → **Subsystem Device Driver DSM**. A command-line interface opens. Enter **`datapath query device`** and press **Enter** to see if there are IBM Flex System V7000 Storage Node disks connected to this host.

Example 12-3 Output from SDDDSM - a disk with four paths is displayed

```
C:\Program Files\IBM\SDDDSM>datapath query device

Total Devices : 1


DEV#:   0  DEVICE NAME: Disk1 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 60050760008F0006F800000000000039
============================================================================
Path#            Adapter/Hard Disk        State  Mode        Select      Errors
    0      Scsi Port3 Bus0/Disk1 Part0     OPEN   NORMAL           0           0
    1      Scsi Port3 Bus0/Disk1 Part0     OPEN   NORMAL         129           0
    2      Scsi Port4 Bus0/Disk1 Part0     OPEN   NORMAL         133           0
    3      Scsi Port4 Bus0/Disk1 Part0     OPEN   NORMAL           0           0

C:\Program Files\IBM\SDDDSM>
```

In the output above note that the UID of the mapped IBM Flex System V7000 Storage Node volume reflects the SERIAL number in the output from **`datapath query device`**.

8. Open the Windows Disk Management window by clicking **Start** → **Run**, typing **`diskmgmt.msc`** or simply open it from the Server Manager. You may have to click **More Actions** → **Rescan Disks** before the FCoE attached disk displays.

   Figure 12-34 shows the Windows Disk Management window with the uninitialized IBM Flex System V7000 Storage Node disk.

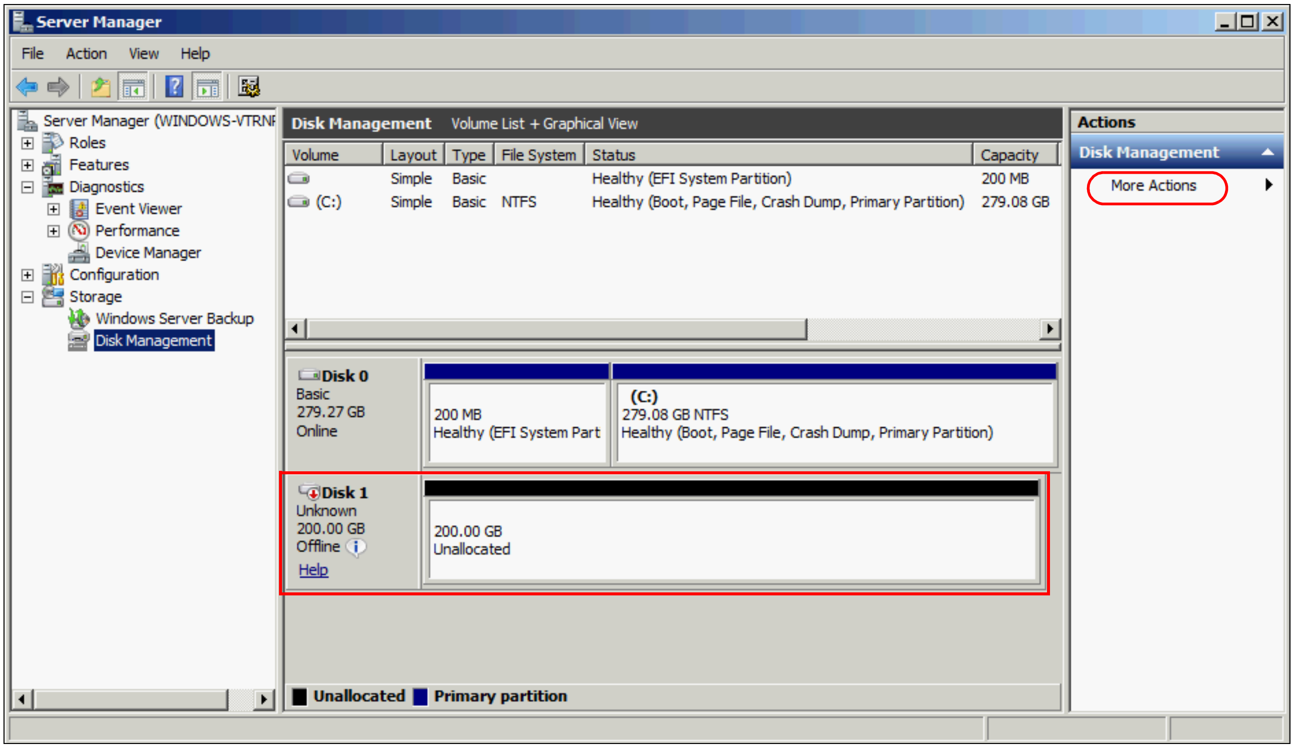


Figure 12-34 Disk Management - FCoE attached disk displays

The next steps for getting access to the Windows disk is to initialize it which was also demonstrated in step 4 on page 478 and forward.

Figure 12-35 on page 498 shows the initialized FCoE attached Windows disk in the Disk Management window.

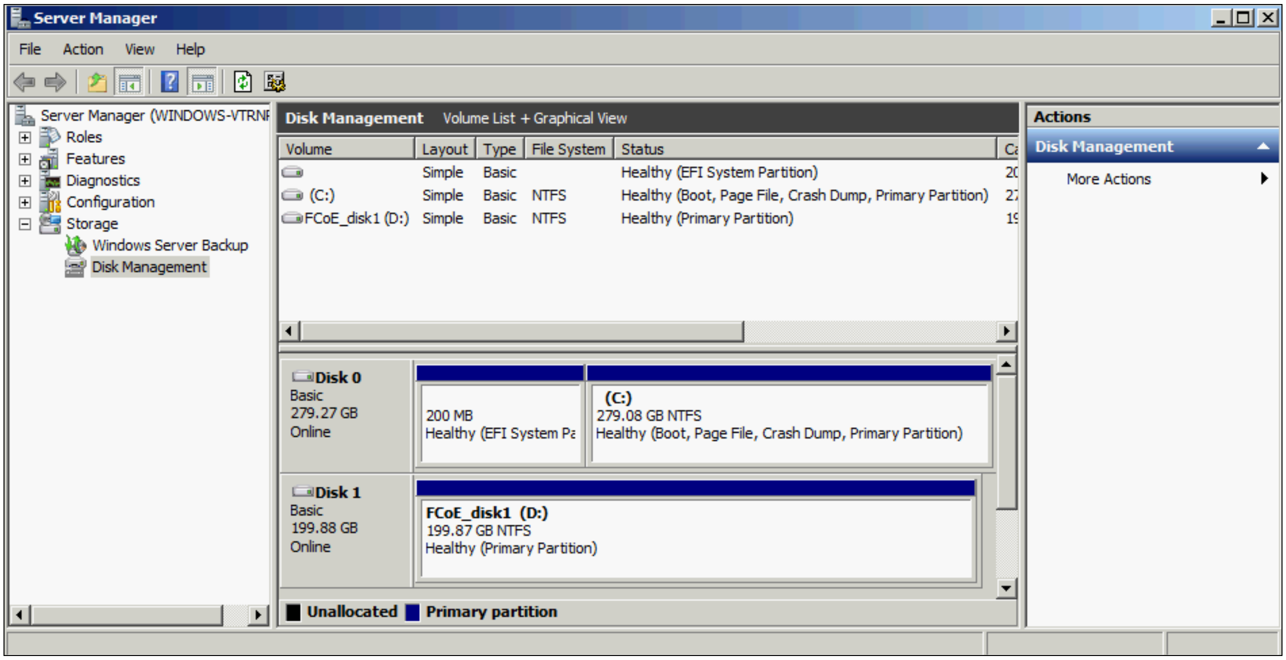


Figure 12-35 Disk Management - FCoE attached disk formatted and online

The setup is now complete and the IBM Flex System V7000 Storage Node is mapping a FCoE attached volume to the server. The Windows 2008 R2 server is prepared to access the volumes over four paths and is able to store data on the storage subsystem.

For more information about FCoE refer to Chapter 11, “SAN Connections and Configuration” on page 431.

12.4 VMware ESX host attachment

In this section we are covering VMware ESX host attachment for Fibre channel (FC) and iSCSI.

12.4.1 VMware ESX Fibre Channel attachment

To perform VMware ESX Fibre Channel attachment, complete the following steps:

1. Right-click your VMware ESX Fibre Channel host in the All Hosts view (Figure 12-36 on page 499) and select **Properties.**

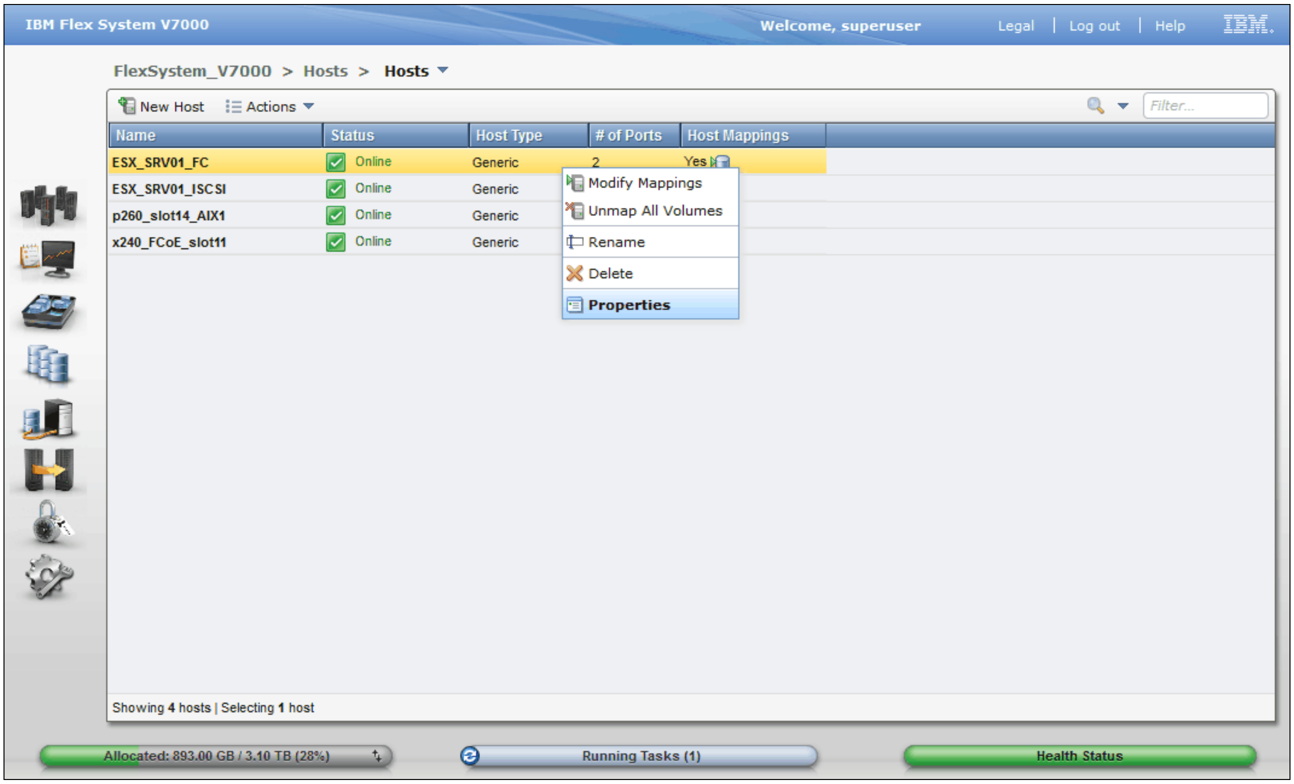


Figure 12-36 Example ESX FC host

Navigate to the **Mapped Volumes** tab (Figure 12-37).

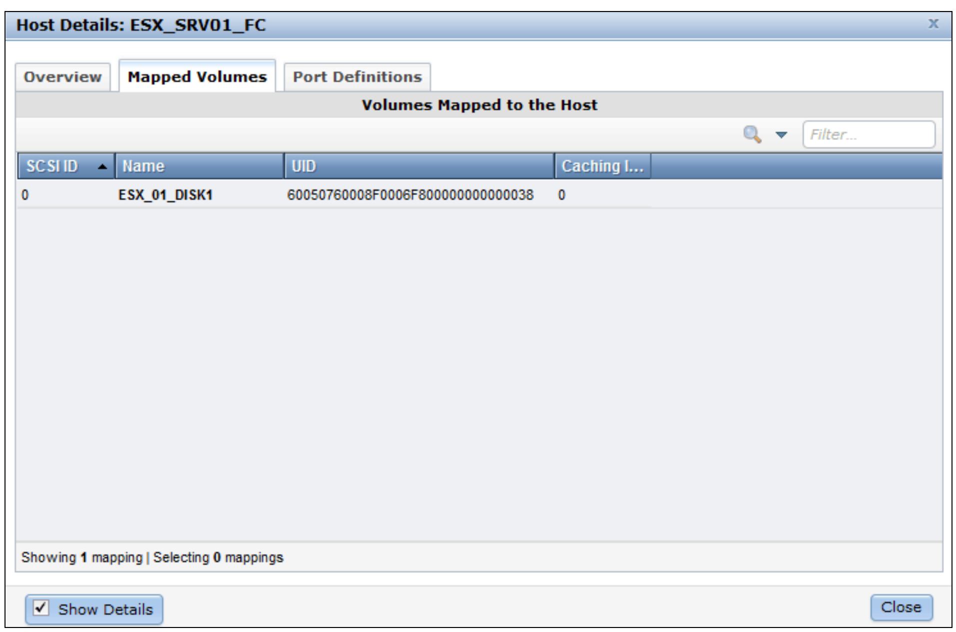


Figure 12-37 Mapped Volumes to ESX FC host

2. In the Host Details window, you see that there is one volume connected to the ESX FC host using SCSI ID 1. The UID of the volume is also displayed.

   Connect to your VMware ESX Server using the vSphere client, navigate to the **Configuration** tab, and select **Storage Adapters** (Figure 12-38).

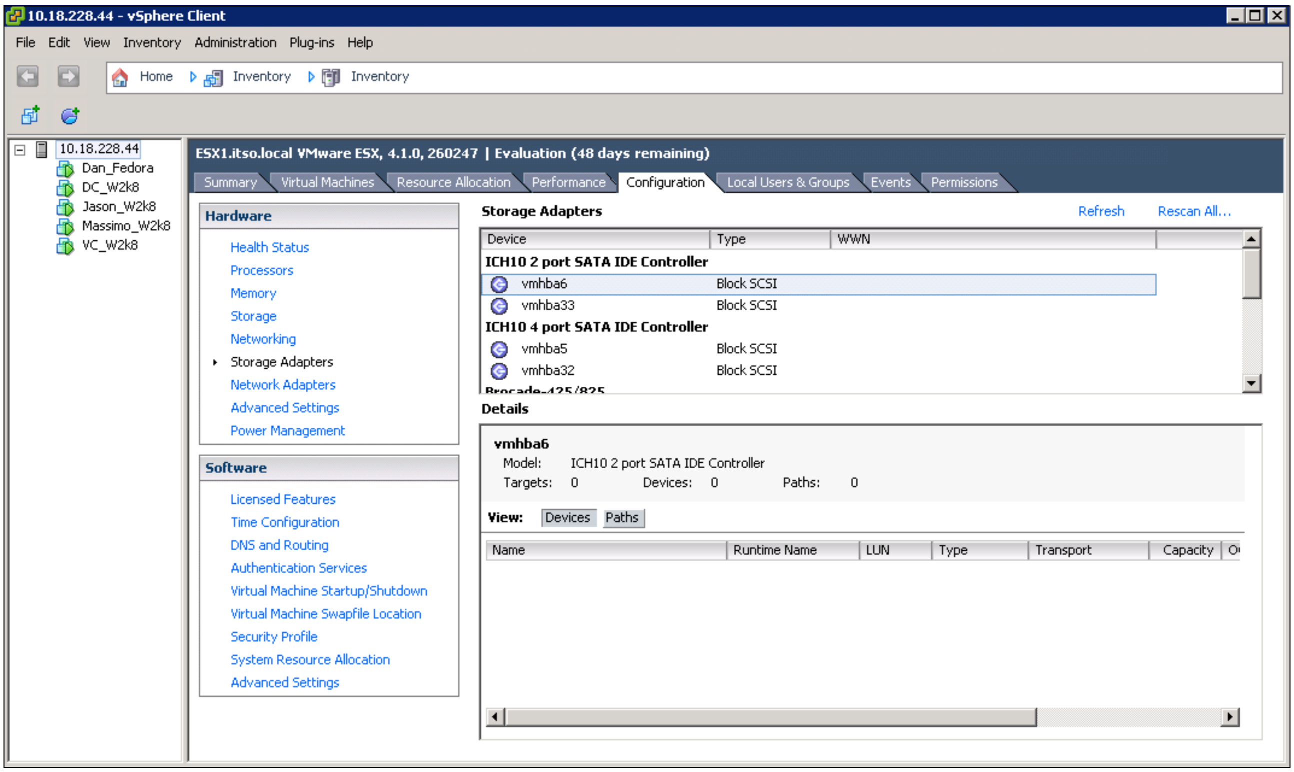


Figure 12-38 vSphere Client: Storage adapters

3. Click **Rescan All...** and click **OK** (Figure 12-39) to scan for new storage devices.

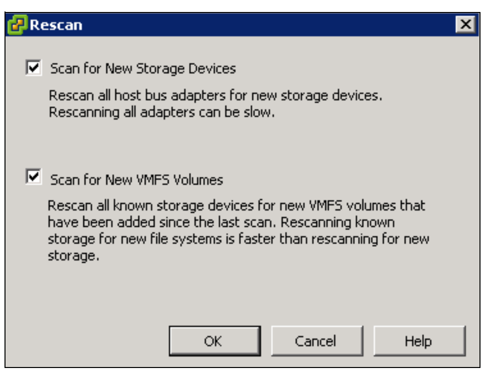


Figure 12-39 Rescan

4. Select **Storage** and click **Add Storage** (Figure 12-40).

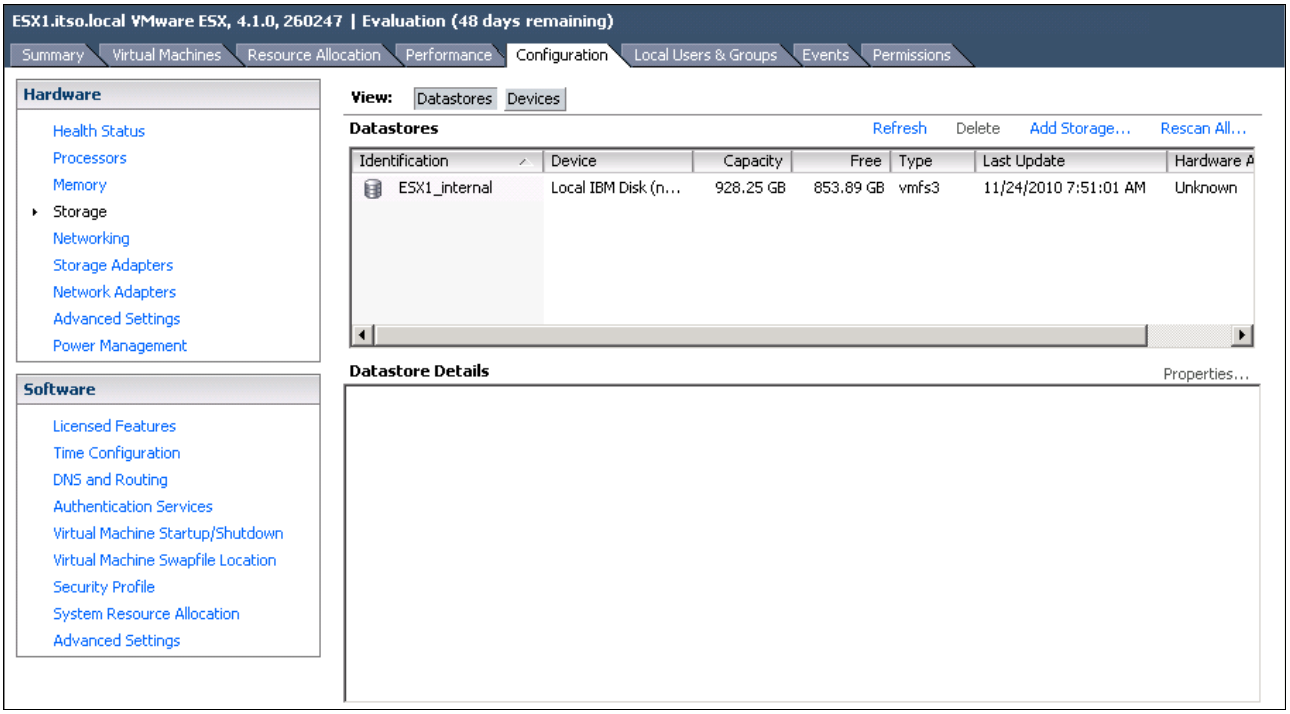


Figure 12-40 vSphere Client: Storage

5. The Add Storage wizard opens. Click **Select Disk/LUN** and click **Next**. The IBM Flex System V7000 Storage Node disk appears (Figure 12-41). Highlight it and click **Next**.

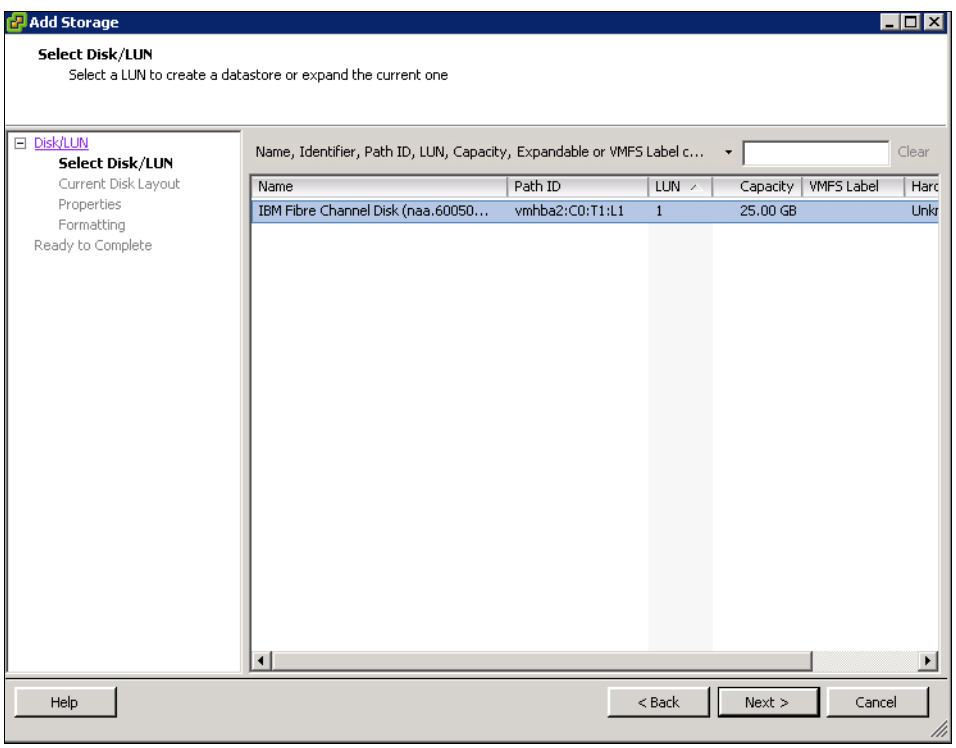


Figure 12-41 Select Disk/LUN menu

6. Follow the wizard to complete the attachment of the disk. After you click **Finish**, the wizard closes and you return to the storage view. In Figure 12-42, you see that the new volume has been added to the configuration.

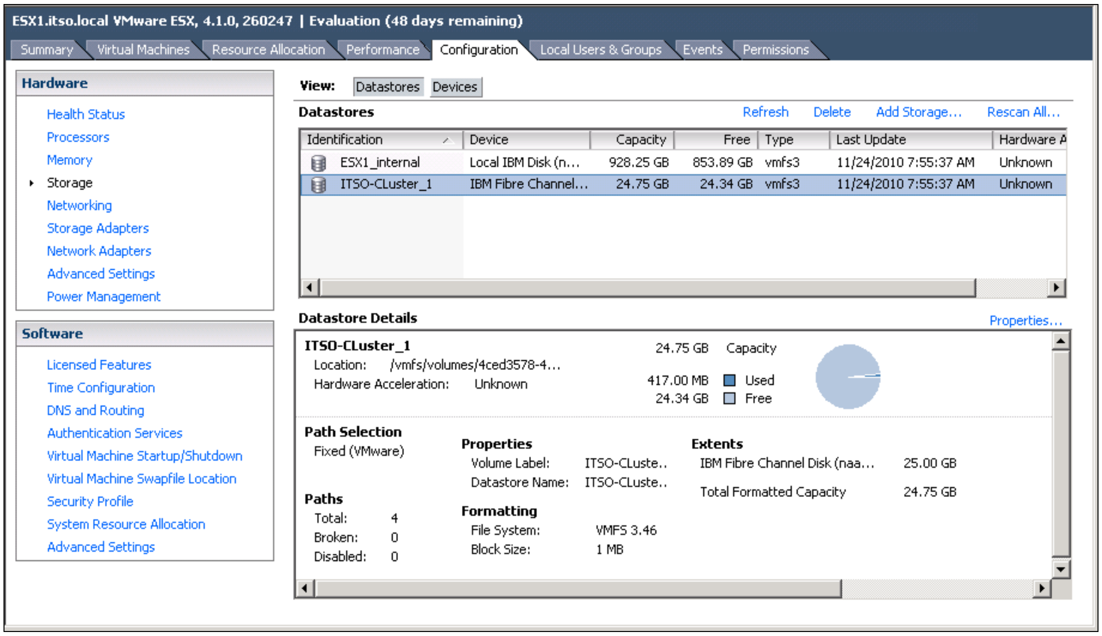


Figure 12-42 Add Storage task complete

7. Highlight the new data store and click **Properties** to see the details of it (Figure 12-43).

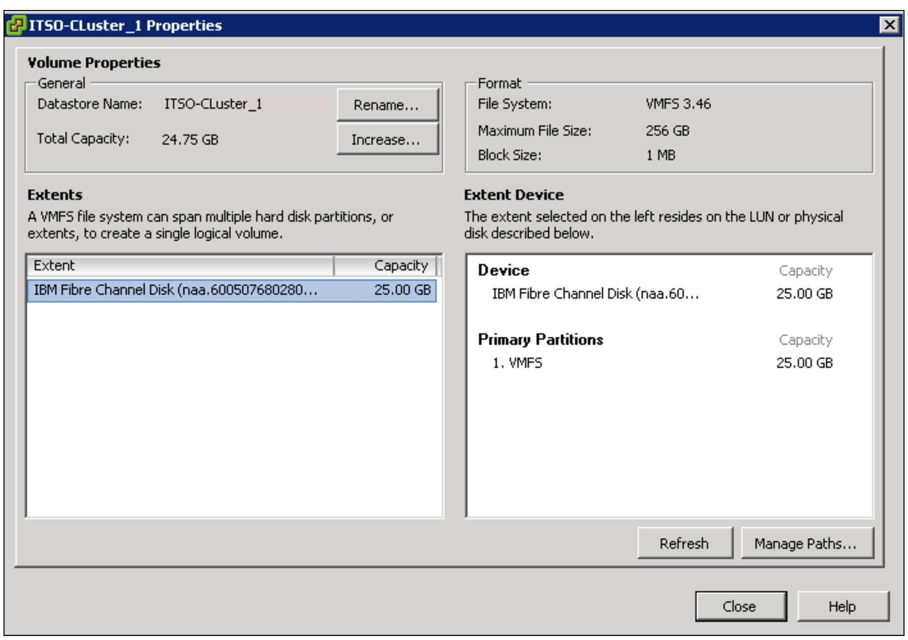


Figure 12-43 Data store properties

8. Click **Manage Paths** to customize the multipath settings. Select **Round Robin** (Figure 12-44) and click **Change**.

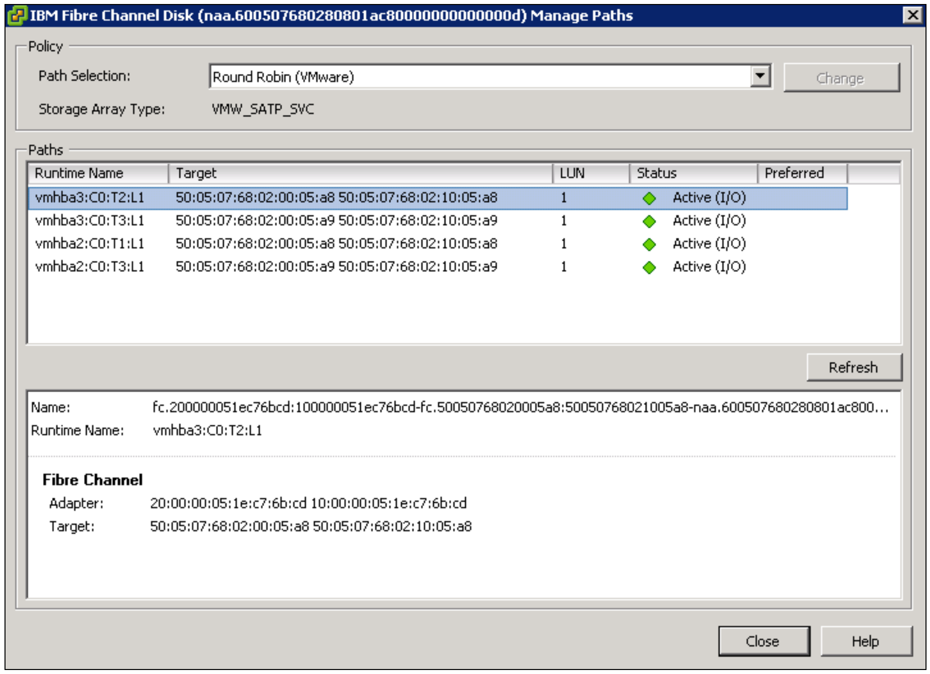


Figure 12-44 Select a data store multipath setting

The storage disk is available and ready to use for your VMware ESX server using Fibre Channel attachment.

12.4.2 VMware ESX iSCSI attachment

To perform a VMware ESX iSCSI attachment, complete the following steps:

1. Right-click your VMware ESX Fibre iSCSI host in the All Hosts view (Figure 12-45) and select **Properties**.

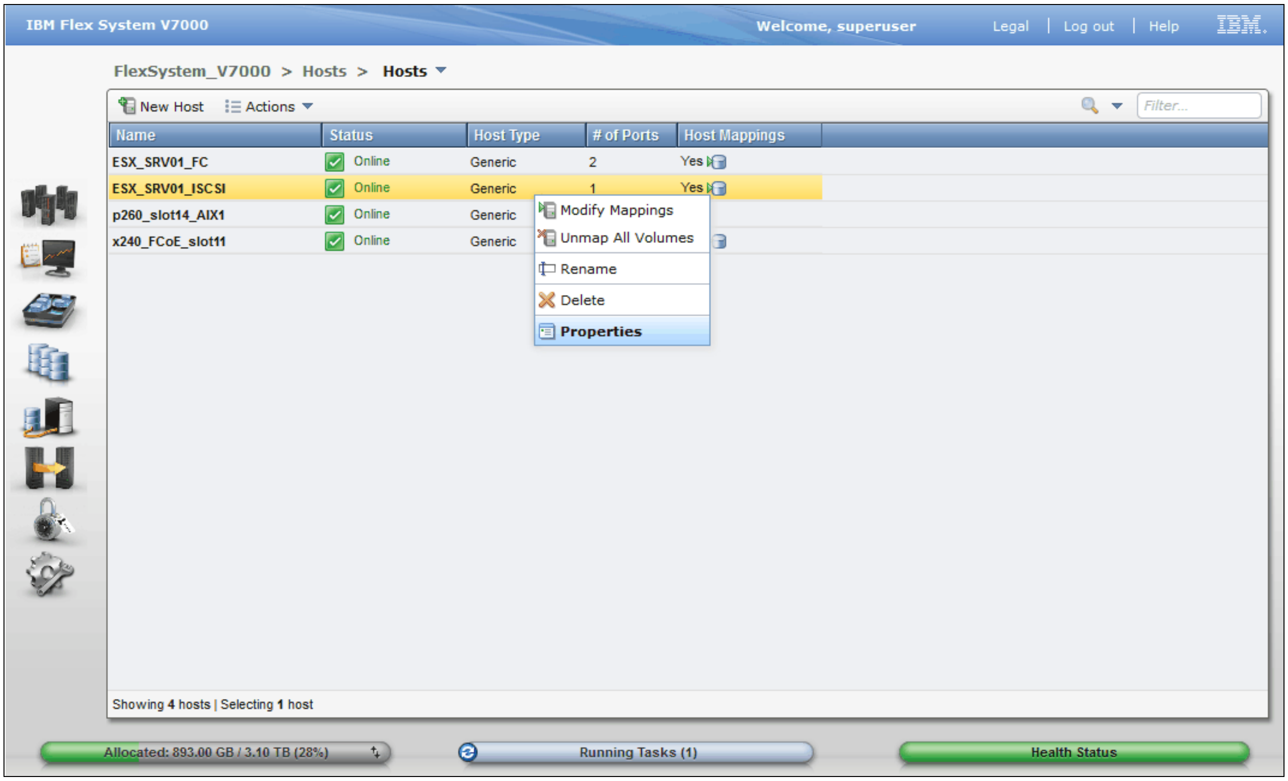


Figure 12-45 Select iSCSI ESX host properties

Navigate to the **Mapped Volumes** tab (Figure 12-46).

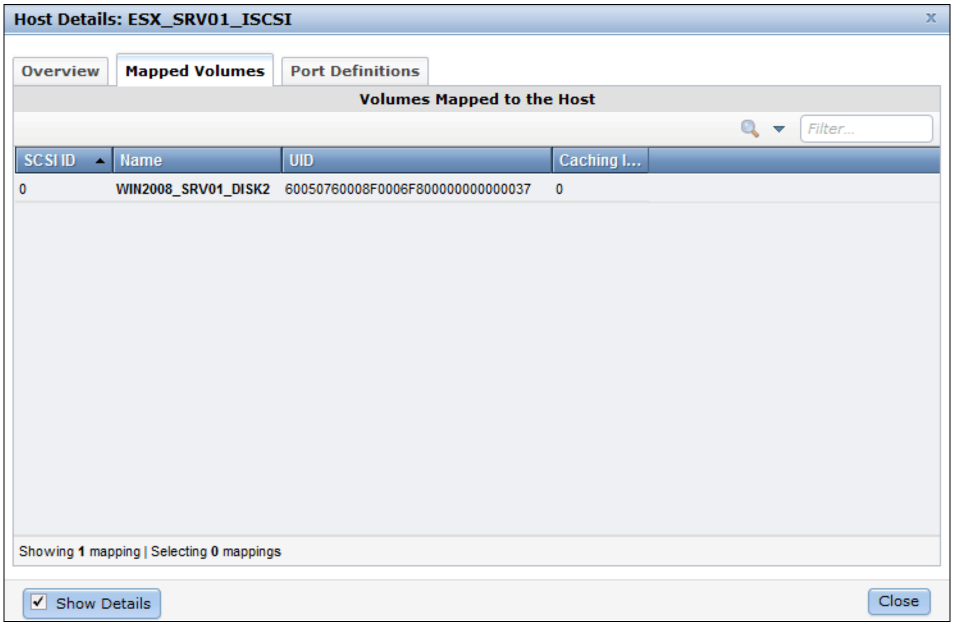


Figure 12-46 iSCSI ESX host properties

2. In the Host Details window, you see that there is one volume connected to the ESX iSCSI host using SCSI ID 0. The UID of the volume is also displayed.

Connect to your VMware ESX Server using the vSphere Client, navigate to the **Configuration** tab, and select **Storage Adapters** (Figure 12-47).

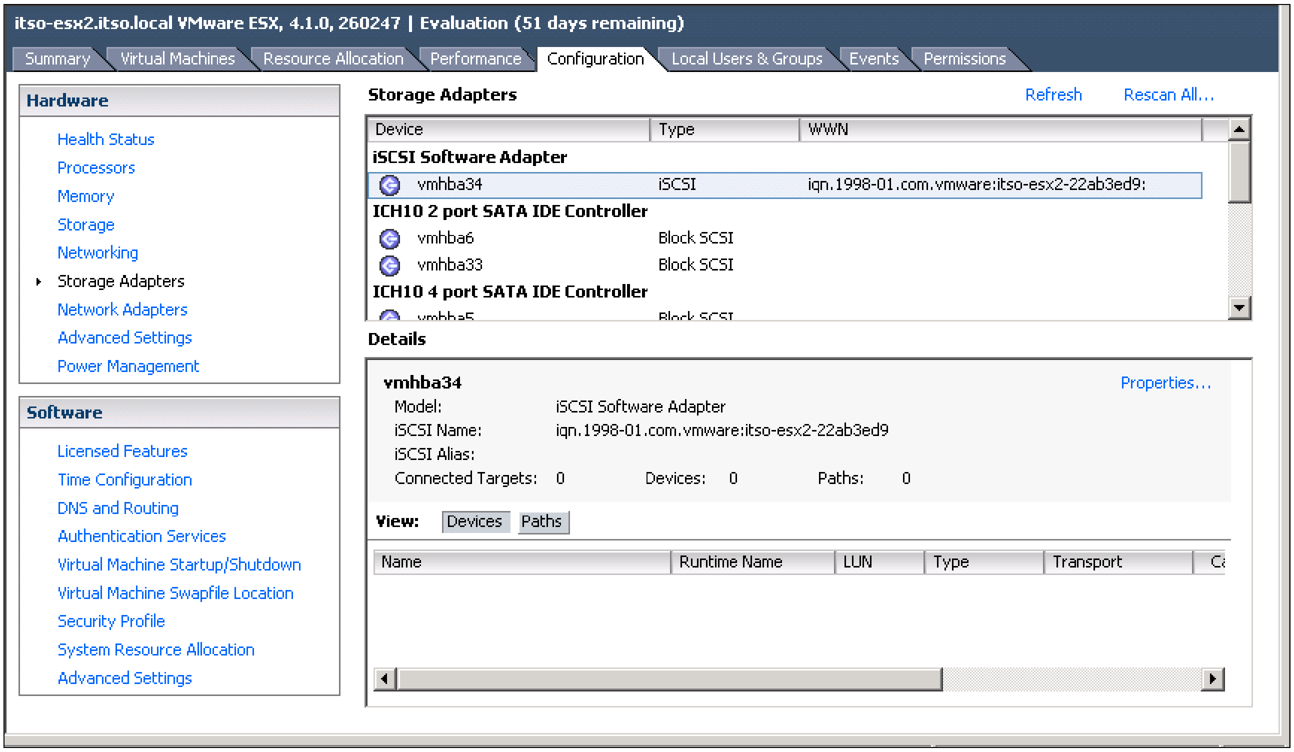


Figure 12-47 vSphere Client: Storage

3. Highlight the **iSCSI Software Initiator** and click **Properties**. The iSCSI initiator properties window opens. Select the **Dynamic Discovery** tab (Figure 12-48) and click **Add**.

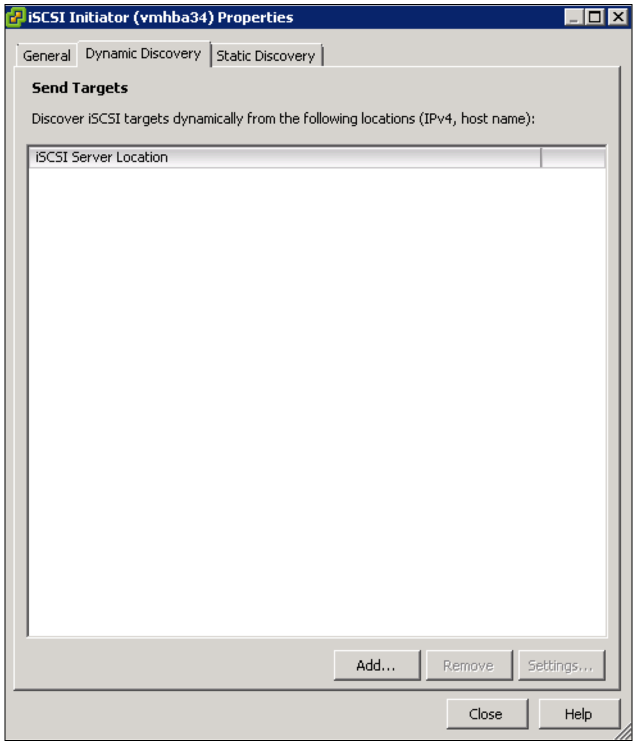


Figure 12-48 iSCSI Initiator properties

4. To add a target, enter the target IP address (Figure 12-49). The target IP address is the IP address of a node in the I/O group from which you are mapping the iSCSI volume. Leave the IP port number at the default value of 3260, and click **OK**. The connection between the initiator and target is established.

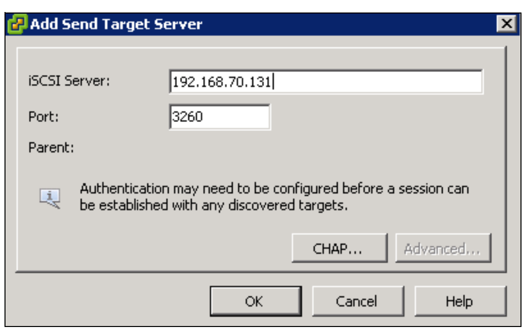


Figure 12-49 Enter a target IP address

Repeat this step for each IBM Flex System V7000 Storage Node iSCSI port you want to use for iSCSI connections.

> **iSCSI IP addresses:** The iSCSI IP addresses are different for the cluster and canister IP addresses; they have been configured in 6.2.2, “Create an iSCSI attached host” on page 259.

5. After you have added all the ports required, close the iSCSI Initiator properties by clicking **Close** (Figure 12-48 on page 507).

   You are prompted to rescan for new storage devices. Confirm the scan by clicking **Yes** (Figure 12-50).

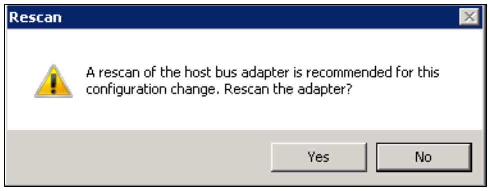


Figure 12-50 Confirm the rescan

6. Go to the storage view shown in Figure 12-51 and click **Add Storage**.

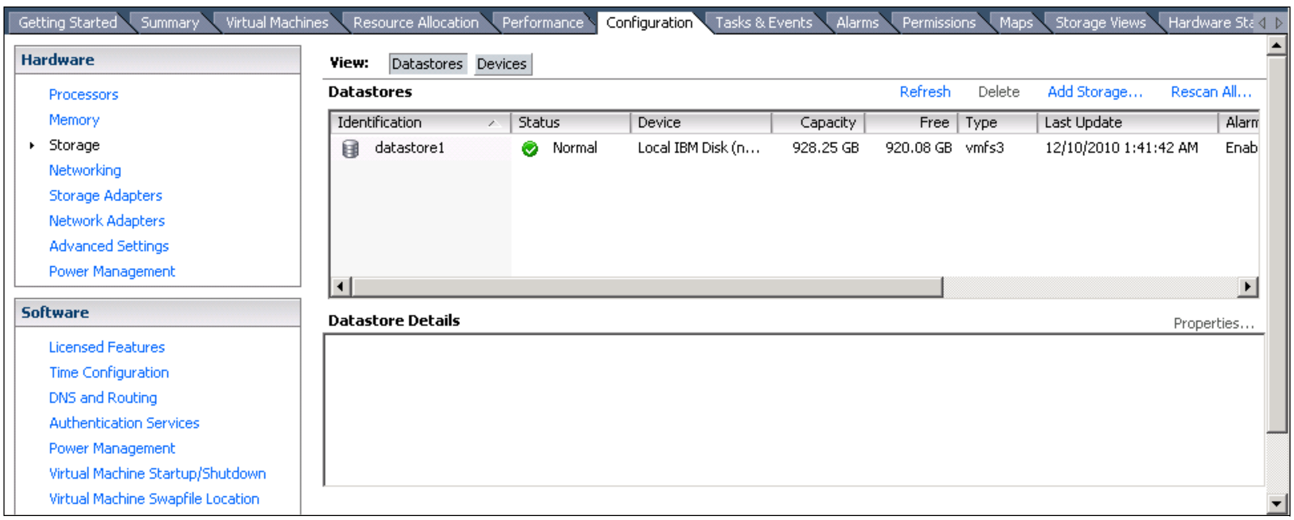


Figure 12-51 Add Storage menu

7. The Add Storage wizard opens (Figure 12-52). Select **Disk/LUN** and click **Next**.

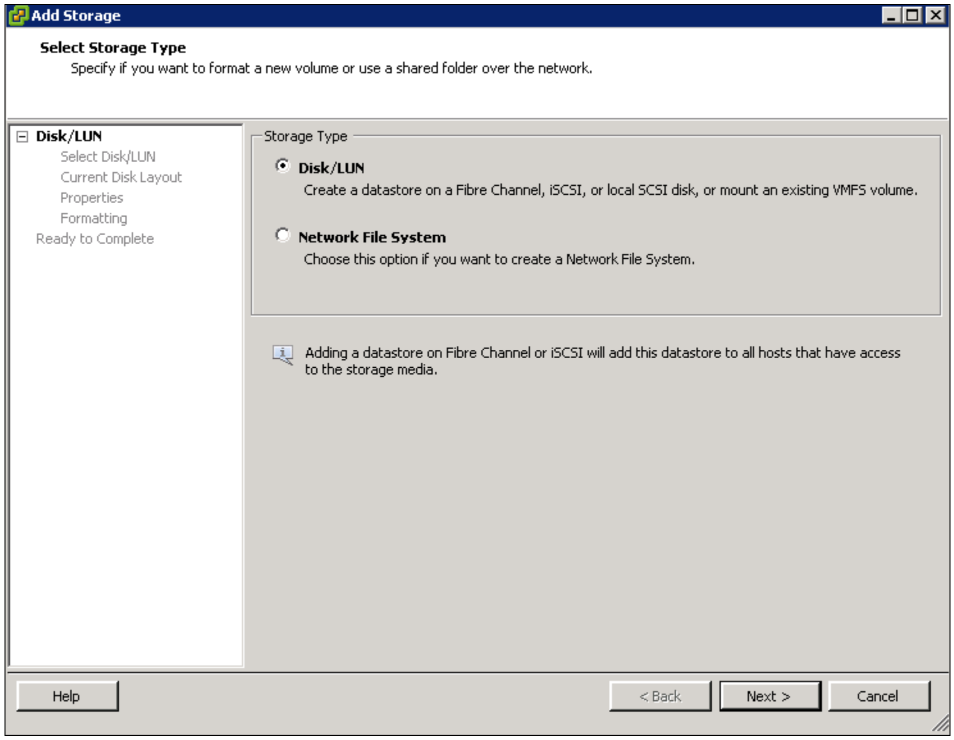


Figure 12-52 Select Disk/LUN menu

8. The new iSCSI LUN displays. Highlight it and click **Next** (Figure 12-53).

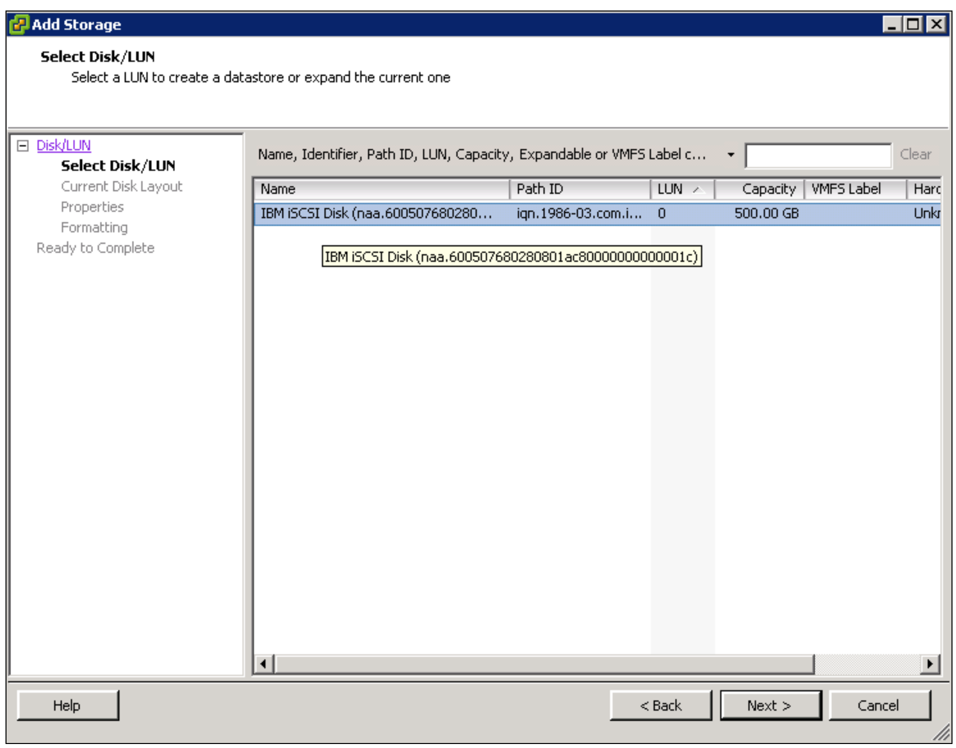


Figure 12-53 Select iSCSI LUN menu

9. Review the disk layout and click **Next** (Figure 12-54).

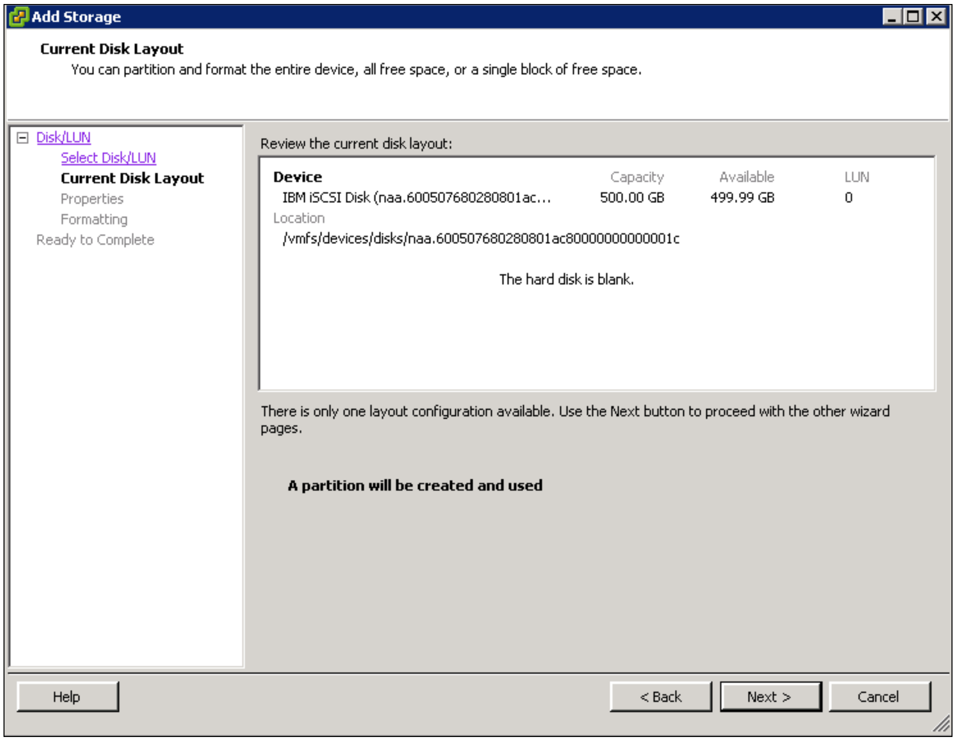


Figure 12-54 Current Disk Layout

10. Enter a name for the data store and click **Next** (Figure 12-55).

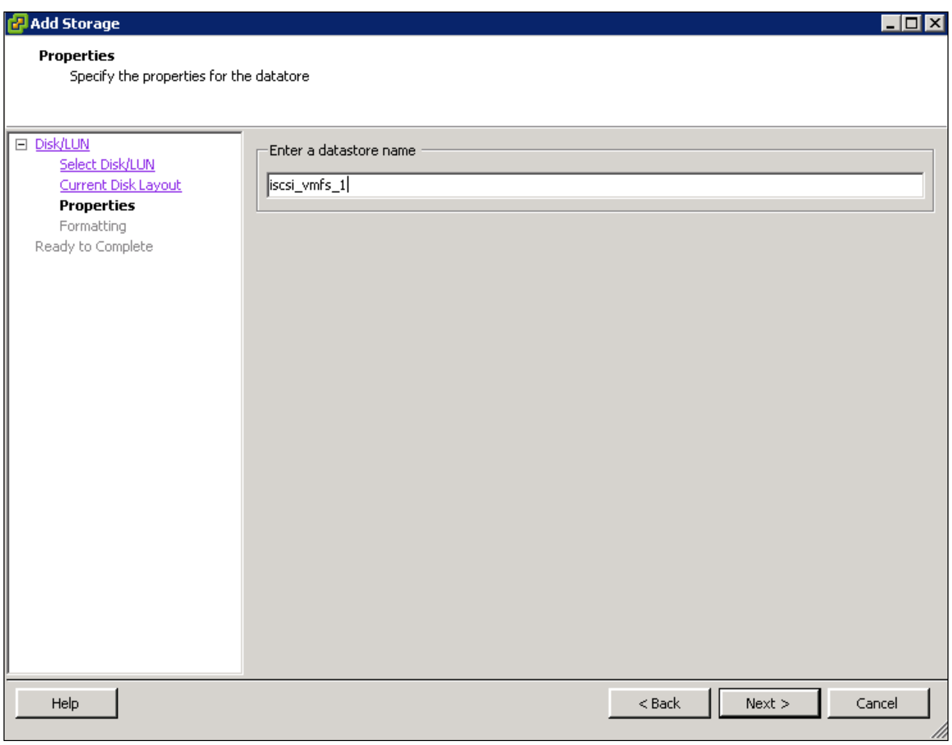


Figure 12-55 Enter a data store name

11.Select the maximum file size and click **Next** (Figure 12-56).

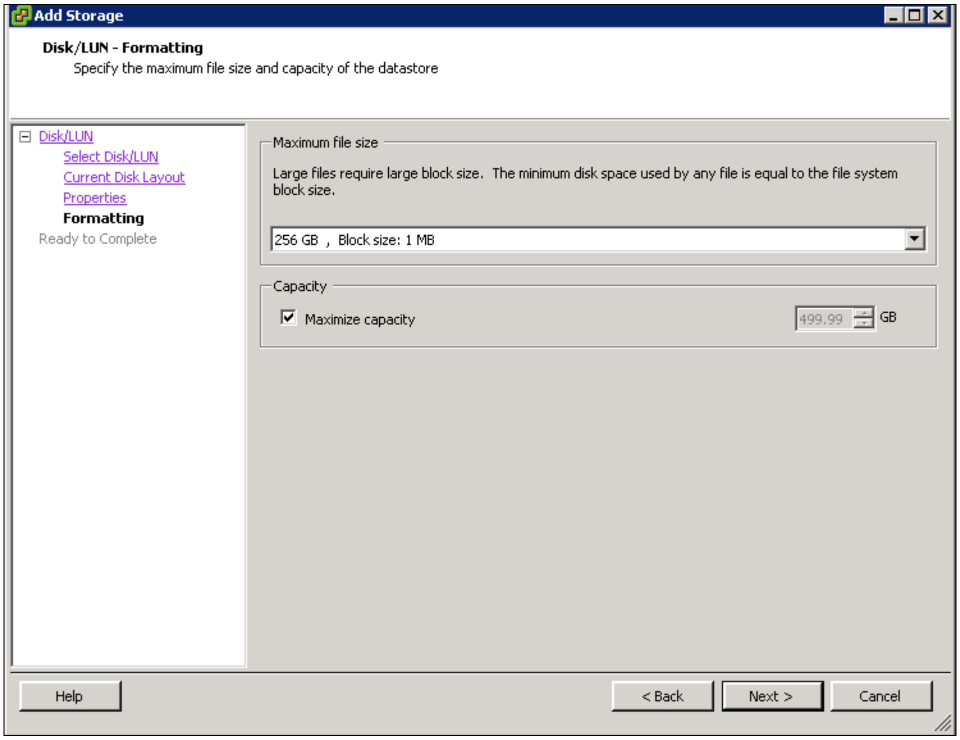


Figure 12-56 Maximum file size

12.Review your selections and click **Finish** (Figure 12-57).

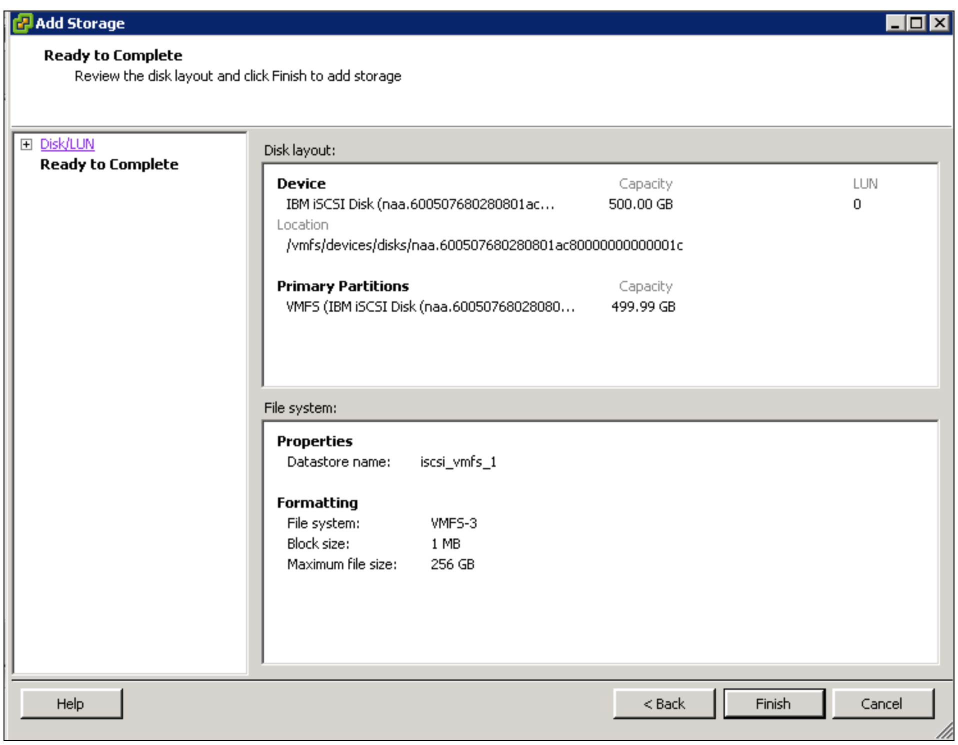


Figure 12-57 Finish the wizard

The new iSCSI LUN is now in the process of being added; this task can take a few minutes. After the tasks complete, the new data store appears in the storage view (Figure 12-58).

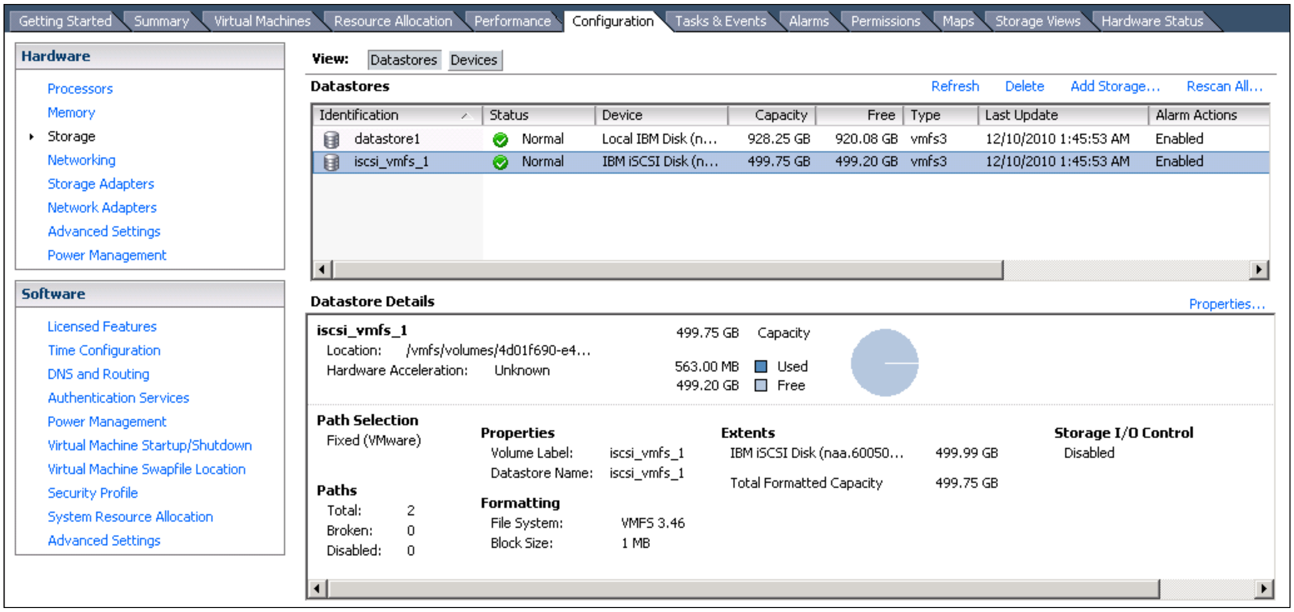


Figure 12-58 New data store available

13. Highlight the new data store and click **Properties** to open and review the data store settings (Figure 12-59).

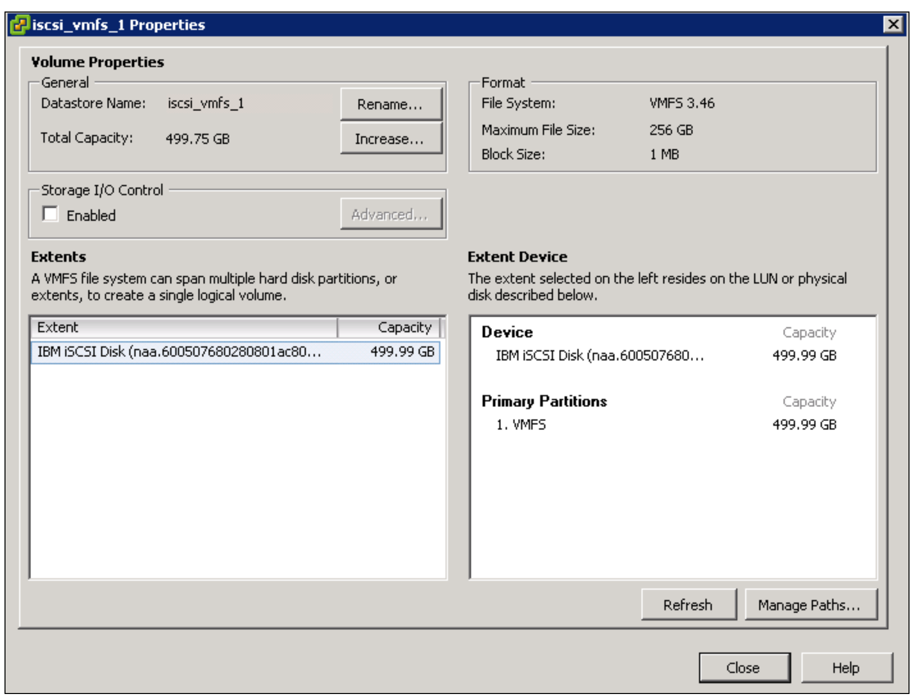


Figure 12-59 Data store properties

14. Click **Manage Paths**, select **Round Robin** as the multipath policy (Figure 12-60), and click **Change**.

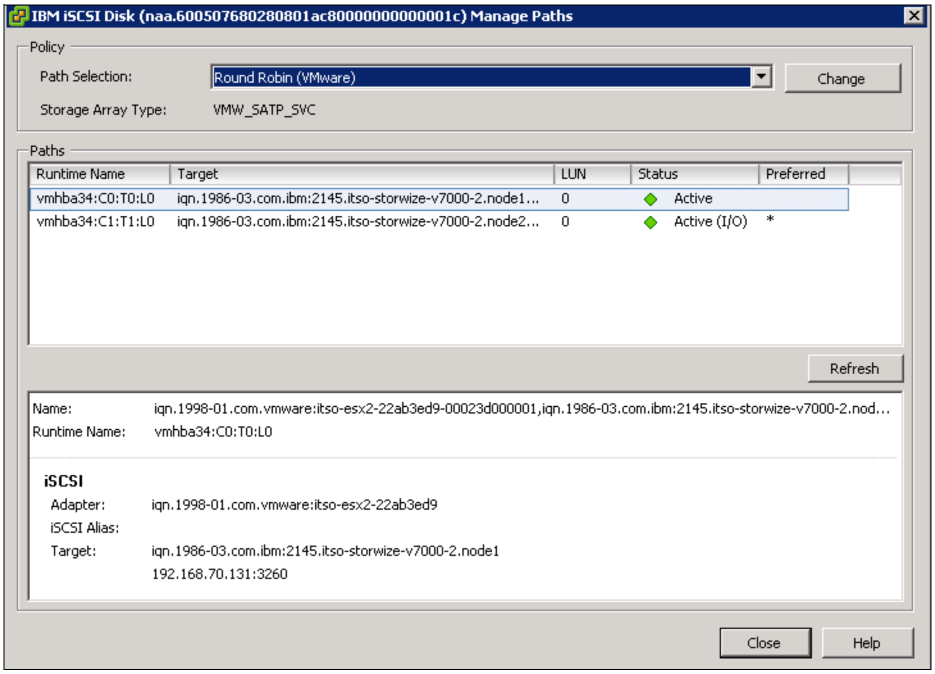


Figure 12-60 Change the multipath policy

15. Click **Close** twice to return to the storage view, and now the storage disk is available and ready to use for your VMware ESX server using an iSCSI attachment.

12.5 AIX Host Attachment

The following section details specific information that relates to the connection of AIX-based compute nodes in the IBM Flex System Enterprise Chassis attaching to the IBM Flex System V7000 Storage Node. The compute node models are the IBM Flex System p260 Compute Node and the IBM Flex System p460 Compute Node. At time of writing FCoE is not yet supported on AIX and we were unable to test iSCSI connectivity. This section detail FC connectivity only which is probably the most popular method of external block level storage connectivity for AIX and VIOS. In the examples below the AIX environment (or VIOS in ***oem_setup_env*** environment) CLI will be used.

12.5.1 Configuring the AIX Compute Node for FC Connectivity

The following list outlines the steps required to attach IBM Flex System V7000 Storage Node volumes to an AIX host, creating these volumes and mapping the volumes to a host compute node have already been addressed in previous chapters.

1. Identify the HBAs in the AIX compute node.
2. Ensure that the correct AIX operating systems and version levels are installed on the compute node, including any updates and Authorized Program Analysis Reports (APARs) on the operating system.
3. Check connectivity of the AIX compute node to the FC switch I/O modules.

4. Check the FC switch zoning.
5. Install the 2145 host attachment support package,
6. Install and configure the Subsystem Device Driver Path Control Module (SDDPCM).
7. Run *cfgmgr* or reboot the compute node to discover and configure the IBM Flex System V7000 Storage Node volumes.

Checking HBAs in the Compute Node

You can display the availability of installed host adapters by using the command ***lsdev -Cc adapter***. In Example 12-4 the HBA adapters are shown as fcs0 and fcs1.

Example 12-4

```
# lsdev -Cc adapter | grep fcs
fcs0    Available 05-00 Dual Port 8Gb FC Mezzanine Card (7710322577107501)
fcs1    Available 05-01 Dual Port 8Gb FC Mezzanine Card (7710322577107501)
```

Check HBA Settings

Some settings on the HBAs differing from the default are recommended, for example, both fast fail and dynamic tracking should be enabled. To check these parameters issue the following command to the FC SCSI I/O Controller Protocol Device of each adapter as show in Example 12-5. From this example we can see that the settings are as they should be.

Example 12-5 Check HBA settings

```
# lsattr -El fscsi0
attach       switch    How this adapter is CONNECTED         False
dyntrk       yes       Dynamic Tracking of FC Devices        True
fc_err_recov fast_fail FC Fabric Event Error RECOVERY Policy True
scsi_id      0x10e00   Adapter SCSI ID                       False
sw_fc_class  3         FC Class for Fabric                   True
# lsattr -El fscsi1
attach       switch    How this adapter is CONNECTED         False
dyntrk       yes       Dynamic Tracking of FC Devices        True
fc_err_recov fast_fail FC Fabric Event Error RECOVERY Policy True
scsi_id      0x10e00   Adapter SCSI ID                       False
sw_fc_class  3         FC Class for Fabric                   True
```

If one of these settings required changing the *chdev* command would be used, for example, to enable fast fail on fcs0 FC SCSI I/O Controller Protocol Device issue command:

chdev -l fscsi0 -a fc_err_recov=fast_fail

12.5.2 Operating system versions and maintenance levels

The following section details the current support information. It is vital that the websites that listed are checked regularly for any updates.

Presently, the only FC HBA supported on the p260 and p460 compute nodes is the IBM Flex System FC3172 2-port 8Gb FC Adapter, refer to SSIC to verify supported configurations:

http://www.ibm.com/systems/support/storage/ssic/interoperability.wss

See IBM ServerProven® website for the latest information about the specific versions and service packs that are supported:

http://ibm.com/servers/eserver/serverproven/compat/us/

For the latest information, and device driver support, *always* refer to the following website:

http://ibm.com/systems/storage/software/virtualization/svc/interop.html

12.5.3 Check Connectivity to IBM Flex System V7000 Storage Node

This is carried out in two stages, firstly connectivity to the I/O Module FC switch, then to ensure that the zoning has been correctly set up we check on the IBM Flex System V7000 Storage Node. For each step we use the WWN of each HBA as a reference. Example 12-6 shows how to display the WWN or 'Network Address' of each adapter.

Example 12-6 Displaying the WWN of each HBA

```
# lscfg -vl fcs0 | grep Network
        Network Address.............21000024FF2F23B4
# lscfg -vl fcs1 | grep Network
        Network Address.............21000024FF2F23B5
```

Switch Connection

The HBAs should be attached to alternate FC switches in the IBM Flex System Enterprise Chassis so both need to be checked. Figure 12-61 shows the switch interface from bay 14 where the compute node resides, the WWN is displayed for HBA fcs0 is shown in the 'Bay14' row. This confirms connectivity to the IBM Flex System FC3171 8Gb SAN Switch. This must be repeated on the second FC switch for the second HBA fcs1.

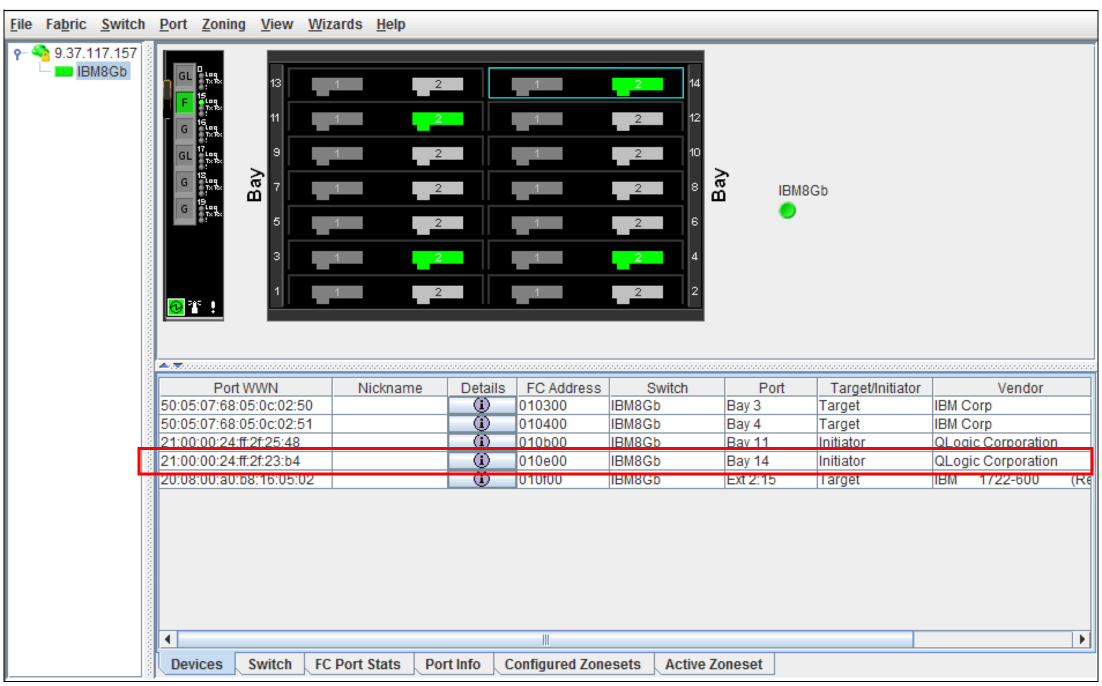


Figure 12-61 Confirm connectivity to the FC switch

Check Zoning

On the IBM Flex System FC3171 8Gb SAN Switch. the zoning is checked as shown in Figure 12-62. Here, the WWN is defined in zone with IBM Flex System V7000 Storage Node canisters.

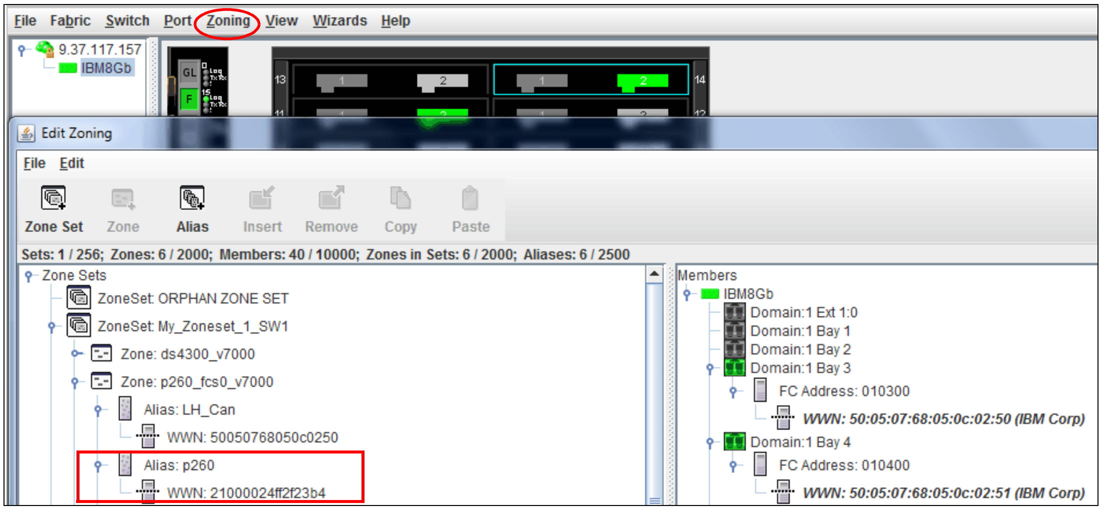


Figure 12-62 Check WWN is defined in zone

Again, this check should be repeated on the second IBM Flex System FC3171 8Gb SAN Switch for the second HBA fcs1.

Check Host Definition on IBM Flex System V7000 Storage Node

From the administration access of IBM Flex System V7000 Storage Node Figure 12-63 shows both HBAs WWNs are correctly displayed in a defined host.

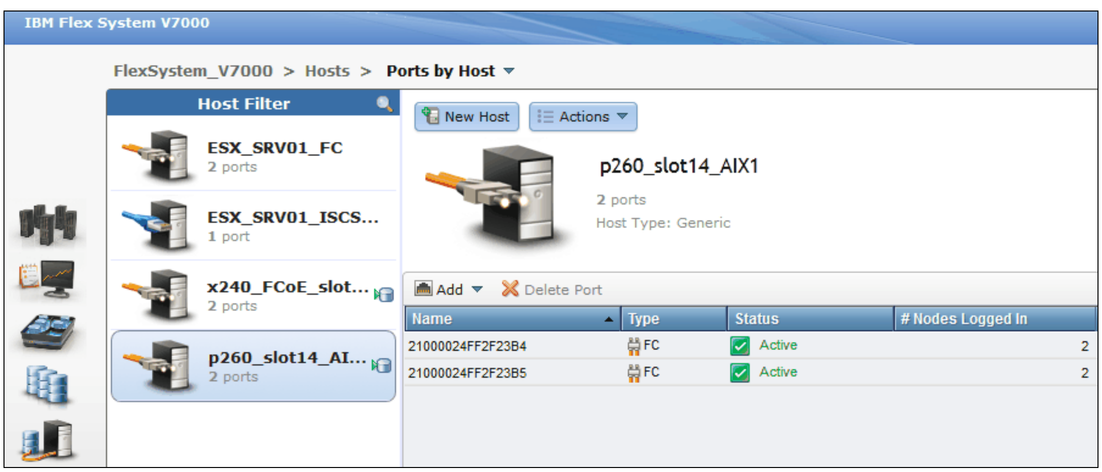


Figure 12-63 WWNs displayed of defined Host

12.5.4 Installing the 2145 host attachment support package

To correctly configure multipathing of IBM Flex System V7000 Storage Node volumes to an AIX compute node requires the proper device type of 2145 which should be displayed when querying with AIX (***lsdev -Cc disk*** command later shown). The 2145 host attachment support fileset is required prior to running ***cfgmgr*** which searches for new devices. Running ***cfgmgr*** prior to installing the host attachment support fileset will result in the LUNs being configured

as “Other SCSI Disk Drives” and will not be recognized by SDDPCM. To correct the device type, hdisks will need to be deleted using ***rmdev -dl hdiskX*** and then ***cfgmgr*** will need to be rerun after the fileset is installed.

Perform the following steps to install the host attachment support package:

1. Access the following website:

http://www.ibm.com/servers/storage/support/software/sdd/downloading.html

2. Search for “Host Attachment for SDDPCM on AIX” in the “Downloads” tab.

3. Download the appropriate host attachment package archive for your AIX version; the fileset contained in the package is *devices.fcp.disk.ibm.mpio.rte* and *devices.sddpcm.71.rte*.

4. Follow the instructions that are provided on the website and the ‘README’ files to install.

12.5.5 Subsystem Device Driver Path Control Module

The Subsystem Device Driver Path Control Module (SDDPCM) is a loadable path control module for supported storage devices to supply path management functions and error recovery algorithms. When the supported storage devices are configured as Multipath I/O (MPIO) devices, SDDPCM is loaded as part of the AIX MPIO FCP (Fibre Channel Protocol) or AIX MPIO SAS (serial-attached SCSI) device driver during the configuration.

The AIX MPIO device driver automatically discovers, configures and makes available all storage device paths. SDDPCM then manages these paths to provide:

- High availability and load balancing of storage I/O
- Automatic path-failover protection
- Concurrent download of supported storage devices licensed machine code
- Prevention of a single-point failure

The AIX MPIO device driver along with SDDPCM enhances the data availability and I/O load balancing of volumes provisioned by the IBM Flex System V7000 Storage Node.

SDDPCM installation

Download the appropriate version of SDDPCM and install using the standard AIX installation procedure as shown in 12.5.4, “Installing the 2145 host attachment support package” on page 516.

Check the driver ‘README’ file and make sure the AIX compute node meets all prerequisites. Example 12-7 on page 518 shows the appropriate version of SDDPCM downloaded into the ***/tmp/sddpcm/lpp*** directory. For our compute node we needed a combination of base level LPPs plus updates. At the beginning of Example 12-7 on page 518 these files are listed. From here we initiate the ***inutoc*** command, this generates a dot toc (.toc) file that is needed by the ***installp*** command prior to installing SDDPCM. We then initiate the ***installp*** command with the “p” preview option to ensure we have everything required.

```
# pwd
/tmp/sddpcm/lpp
# ls
devices.fcp.disk.ibm.mpio.rte  devices.sddpcm.71.rte
devices.sddpcm.71.2.6.3.2.bff
# inutoc .
# installp -acgp -d . devices.sddpcm.71.rte > preview.txt 2>&1
```

```
# ls -al
total 8648
drwxr-xr-x    2 root     system          256 Nov  1 21:14 .
drwxr-xr-x    4 root     system         4096 Nov  1 21:04 ..
-rw-r--r--    1 root     system         2455 Nov  1 21:12 .toc
-rw-------    1 root     system        95232 Nov  1 21:08 devices.fcp.disk.ibm.m
pio.rte
-rw-r--r--    1 root     system      1188864 Nov  1 21:07 devices.sddpcm.71.2.6.
3.2.bff
-rw-r--r--    1 root     system      3123200 Nov  1 21:07 devices.sddpcm.71.rte
-rw-r--r--    1 root     system         1852 Nov  1 21:14 preview.txt
```

The output file '*preview.text*' is checked to ensure all prerequisite LPPs exist and SDDPCM will install. If there are no errors then we can remove the preview option and run the installp command again to install the LPPs.

Example 12-7 shows ***installp*** command and the ***lslpp*** command that can be used to check the version of SDDPCM is successfully installed.

Example 12-7

```
# installp -acg -d . devices.sddpcm.71.rte > install.txt 2>&1
# lslpp -l  devices.sddpcm.71.rte
  Fileset                      Level  State      Description
  ----------------------------------------------------------------------------
Path: /usr/lib/objrepos
  devices.sddpcm.71.rte      2.6.3.2  COMMITTED  IBM SDD PCM for AIX V71

Path: /etc/objrepos
  devices.sddpcm.71.rte      2.6.3.2  COMMITTED  IBM SDD PCM for AIX V71
```

The machine will require a reboot following the LPP install, this can be verified by messages in the '*install.txt'* output file created in Example 12-7.

Note: If SDDPCM is installed prior to IBM Flex System V7000 Storage Node volumes being allocated then the ***cfgmgr*** command must be run so that AIX will detect the newly allocated volumes with the correct SDDPCM drivers.

Checking and Configuring assigned volumes with AIX

We use an AIX host with host name p260_slot14_AIX as shown in Figure 12-63 on page 516 to show the IBM Flex System V7000 Storage Node defined with volumes mapped. Example 12-8 shows host configuration after the machine has rebooted. The ***lspv*** output shows existing and new hdisks and ***lsvg*** output shows existing Volume Group.

Example 12-8 New volumes detected by AIX

```
# lspv
hdisk0          000f857a4113adda                    rootvg          active
hdisk1          none                                None
hdisk2          none                                None
# lsvg
rootvg
# lsdev -Cc disk
hdisk0 Available 01-00-00 SAS Disk Drive
```

```
hdisk1 Available 05-00-01 MPIO FC 2145
hdisk2 Available 05-00-01 MPIO FC 2145
```

The ***lsdev*** command confirms that the new drives hdisk1 and hdisk2 are FC 2145 (IBM Flex System V7000 Storage Node volumes) hdisks.

The ***mkvg*** command can now be used to create a Volume Group with the two newly configured hdisks. We then use the new hdisks to create filesystems and make them available for users as shown in Example 12-9.

Example 12-9 Making newly attached storage available for use.

```
# mkvg -y itsoaixvg1 hdisk1 hdisk2
0516-1254 mkvg: Changing the PVID in the ODM.
0516-1254 mkvg: Changing the PVID in the ODM.
itsoaixvg1
# mklv -a c -t jfs2 -y test1lv itsoaixvg1 200 hdisk1
test1lv
# mklv -a c -t jfs2 -y test2lv itsoaixvg1 100 hdisk2
test2lv
# crfs -v jfs2 -a logname=INLINE -A y -p rw -d test1lv -u v7kgrp -m /test1
File system created successfully.
13055396 kilobytes total disk space.
New File System size is 26214400
# crfs -v jfs2 -a logname=INLINE -A y -p rw -d test2lv -u v7kgrp -m /test2
File system created successfully.
6527596 kilobytes total disk space.
New File System size is 13107200
# mount -t v7kgrp
```

The ***lspv*** command output now shows the new Volume Group label on each of the hdisks that were included in the Volume Groups. Also, new filesystems are now made available for use as seen in Example 12-10

Example 12-10 New Volume Group and Filesystems available for use.

```
# lspv
hdisk0          000f857a4113adda                    rootvg          active
hdisk1          000f857abf0e85c2                    itsoaixvg1      active
hdisk2          000f857abf0e8761                    itsoaixvg1      active
# df -m
Filesystem    MB blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4        1024.00    904.55   12%     4040     2% /
/dev/hd2        2048.00    620.31   70%    30215    18% /usr
/dev/hd9var     1024.00   1002.39    3%      432     1% /var
/dev/hd3        3072.00   2182.62   29%     1249     1% /tmp
/dev/fwdump     1536.00   1511.81    2%        9     1% /var/adm/ras/platform
/dev/hd1         512.00    511.59    1%        5     1% /home
/dev/hd11admin     512.00    511.59    1%        5     1% /admin
/proc                 -         -    -         -     -  /proc
/dev/hd10opt     512.00    494.11    4%      235     1% /opt
/dev/livedump    512.00    511.42    1%        5     1% /var/adm/ras/livedump
/dev/test1lv   12800.00  12747.73    1%        4     1% /test1
/dev/test2lv    6400.00   6373.70    1%        4     1% /test2
```

Using SDDPCM

The SDDPCM is administered using the ***pcmpath*** command. This commands is used to perform all administrative functions such as displaying and changing the path state. The ***pcmpath query adapter*** command displays the current state of the adapters. In Example 12-11, we can see the adapter status and that both adapters are showing as optimal with State=NORMAL and Mode=ACTIVE.

Example 12-11 SDDPCM commands that are used to check the availability of the adapters

```
# pcmpath query adapter

Total Dual Active and Active/Asymmetric Adapters : 2

Adpt#     Name     State      Mode              Select      Errors  Paths  Active
    0   fscsi0    NORMAL    ACTIVE                 949           0      4       4
    1   fscsi1    NORMAL    ACTIVE                 991           0      4       4
```

The ***pcmpath query device*** command displays the current state of the volume devices and their paths. In Example 12-12, we can see the path State and Mode for each of the defined hdisks. The status that both adapters are showing as optimal with State=OPEN and Mode=NORMAL.

Additionally, an asterisk (*) displayed next to paths indicates inactive paths that are configured to the non-preferred IBM Flex System V7000 Storage Node canisters in the IO Group.

Example 12-12 SDDPCM commands that are used to check the availability of the devices

```
# pcmpath query device

Total Dual Active and Active/Asymmetric Devices : 2


DEV#:   1  DEVICE NAME: hdisk1  TYPE: 2145  ALGORITHM:  Load Balance
SERIAL: 60050760008F0006F800000000000029
==========================================================================
Path#      Adapter/Path Name          State     Mode      Select     Errors
    0           fscsi0/path0           OPEN   NORMAL         574          0
    1*          fscsi0/path1           OPEN   NORMAL          42          0
    2           fscsi1/path2           OPEN   NORMAL         633          0
    3*          fscsi1/path3           OPEN   NORMAL          42          0

DEV#:   2  DEVICE NAME: hdisk2  TYPE: 2145  ALGORITHM:  Load Balance
SERIAL: 60050760008F0006F80000000000002A
==========================================================================
Path#      Adapter/Path Name          State     Mode      Select     Errors
    0*          fscsi0/path0           OPEN   NORMAL          42          0
    1           fscsi0/path1           OPEN   NORMAL         293          0
    2*          fscsi1/path2           OPEN   NORMAL          42          0
    3           fscsi1/path3           OPEN   NORMAL         276          0
#
```

12.6 Linux host attachment

This section briefly guide you in preparing and configuring a Linux host for FC attachment.

12.6.1 Linux Fibre Channel attachment

Follow these steps to configure the Linux host:

- Use the latest firmware levels on your host system.
- Install the HBA or HBAs on the Linux server.
- Install the supported HBA driver/firmware and upgrade the kernel, if required.
- Connect the Linux server FC host adapters to the switches.
- Configure the switches (zoning) if needed.
- Install SDD for Linux.
- Configure the host, volumes, and host mapping in the IBM Flex System V7000 Storage Node.
- Rescan for LUNs on the Linux server to discover the volumes that were created on the IBM Flex System V7000 Storage Node.

The IBM Flex System V7000 Storage Node supports hosts that run the following Linux distributions:

- Red Hat Enterprise Linux
- SUSE Linux Enterprise Server

For more information connecting a Linux host to the IBM Flex System V7000 Storage Node, refer to *Implementing the IBM System Storage SAN Volume Controller V6.3*, SG24-7933

13

Maintenance and Troubleshooting

There are many ways to monitor and troubleshoot the IBM Flex System V7000 Storage Node. This chapter describes the following topics:

- Reliability, availability, and serviceability (RAS)
- Hardware and LEDs descriptions
- Monitoring system status and health
- Managing IBM Flex System V7000 Storage Node using FSM
- Configuration backup and restore process
- Software upgrade procedure
- Drive Firmware upgrade procedure
- Troubleshooting
- Audit log navigation
- Support data collection
- Using event notification
- Configuring Call Home
- Event log navigation
- Shutting down and power on Storage Node procedure

For more information on how to service, maintain, and troubleshoot the Flex System V7000 Storage Node, refer to IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide (GC27-4205) available at :

http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_bkmap_trblshoot.pdf

13.1 Reliability, availability, and serviceability (RAS)

Reliability, availability, and serviceability (RAS) are important concepts in the design of the IBM Flex System V7000 Storage Node. Hardware features, software features, design considerations, and operational guidelines all contribute to make the IBM Flex System V7000 Storage Node reliable.

Fault tolerance and high levels of availability are achieved by:

- The RAID capabilities of the underlying disk subsystems
- IBM Flex System Flex System V7000 nodes clustering using a *Compass* architecture
- Auto-restart of hung nodes
- Battery units to provide memory protection in the event of a site power failure
- Host system failover capabilities

High levels of serviceability are achieved by providing:

- Cluster error logging
- Asynchronous error notification
- Dump capabilities to capture software detected failures
- Concurrent diagnostic procedures
- Directed maintenance procedures
- Concurrent log analysis and memory dump data recovery tools
- Concurrent maintenance of all IBM Flex System V7000 components
- Concurrent upgrade of IBM Flex System V7000 Software and microcode
- Concurrent addition or deletion of the IBM Flex System V7000 nodes canister in a cluster
- Software recovery through a service panel push button
- Automatic software version correction when replacing a node
- Detailed status and error conditions displayed on the service panel
- Error and event notification through SNMP, syslog, and email

At the heart of the IBM Flex System V7000 Storage Node is a pair of node canisters. These two canisters share the data transmitting and receiving load between the attached hosts and the disk arrays. This section looks at the RAS features of the IBM Flex System V7000 Storage Node system monitoring, and troubleshooting.

13.2 Hardware and LEDs descriptions

This documentation helps you determine the system state using the LED indicators on the enclosure

13.2.1 Understanding the system state using the control enclosure LEDs

The IBM Flex System V7000 Storage Node control enclosure is installed in the IBM Flex System Enterprise Chassis and the machine type and model (MTM) is as followes:

- Machine type and model (MTM) for the control enclosure is 4939-A49, 4939-H49, or 4939-X49.
- Machine type and model (MTM) for the expansion enclosure is 4939-A29, 4939-H29, or 4939-X29.

It contains two node canisters and up to 24 2.5-inch drives. Both spinning hard disk drives (HDD) and solid state drives (SSD) are supported.

Note: Be aware of the differences between the 4939 canisters and the enclosures, you must be able to distinguish between them when you service the system. To obtain the machine type, model number, and serial number from an IBM Flex System V7000 Storage Node enclosure, find a set of blue pull-out tabs on the front of the 4939 enclosure, as shown in Figure 13-1

You must know the correct serial number and MTM when you contact IBM Support.

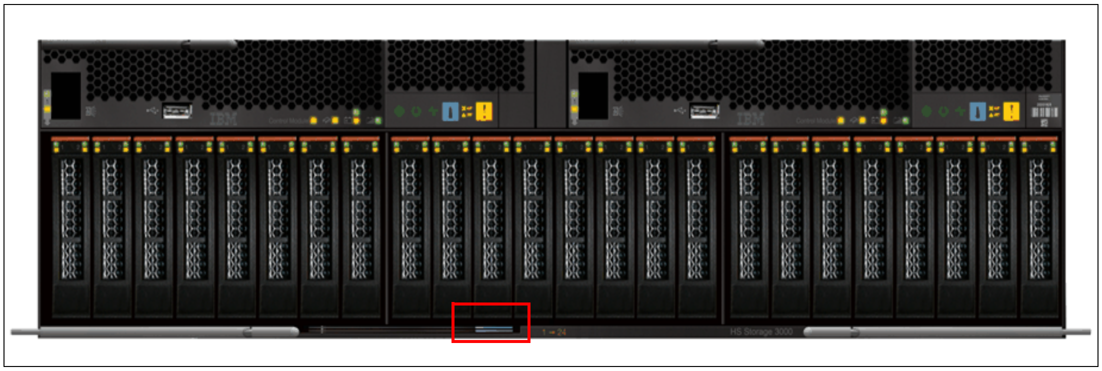


Figure 13-1 Locate MTM info from a an IBM Flex System V7000 Storage Node enclosure

The LEDs provide a general idea of the state of the system. You can obtain more detail from the IBM Flex System V7000 Storage Node management GUI and the Service Assistant. Examine the LEDs when you are not able to access the management GUI or the Service Assistant, or when the system is not showing any information about a device.

The first step is to Identify and locate the control enclosure for the system that you are troubleshooting. Next is to determine the state of the control enclosure, which includes the node canisters with a battery in each canister. Your control enclosure is operational if you can manage the system using the management GUI.

Table 13-1 defines the various LED states for the node canisters in the IBM Flex System V7000 Storage Node.

Table 13-1 Definition of a node canister LED states

State	Description
Off	The LED is continuously not lit or off
Blinking slowly	The LED turns on and off at a frequency of 1 Hz: It is on for 500 ms, then off for 500 ms, then repeats.
Blinking	The LED turns on and off at a frequency of 2 Hz: It is on for 250 ms, then off for 250 ms, then repeats.
Blinking fast	The LED turns on and off at a frequency of 4 Hz: It is on for 125 ms, then off for 125 ms, then repeats.
On	The LED is continuously lit or on
Flashing	The LED is lit to indicate some activity, then turns off. The rate and duration that the LED is lit depends on the rate and duration of the activity.

IBM Flex System V7000 Storage Node has indicators or light-emitting diodes (LEDs) that provide status of the various components to help you determine the system state and

appropriate actions to take to resolve the issues. See Figure 13-2 for a description of the various LED states that you might observe.

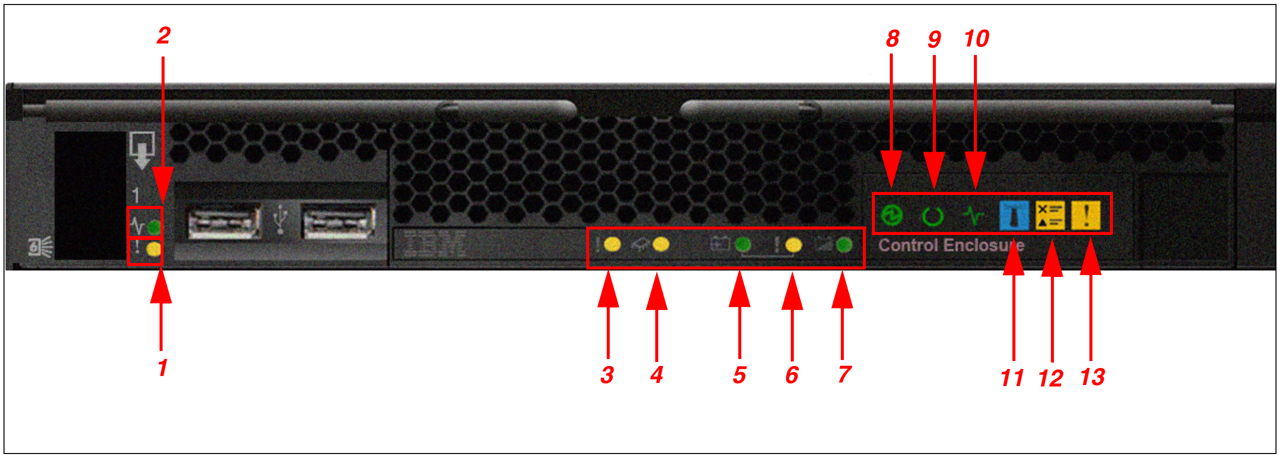


Figure 13-2 Node canister LED layout

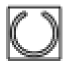
1. Start by considering the state of the power LED **8**, as shown in Figure 13-2 and Table 13-2. Continue to troubleshoot the canister until the power LED is ON.

Table 13-2 Node canister Power LED states

LED layout	LED Name	Color	States
8	 Power	Green	► **OFF** – There is no power to the canister. Make sure that the CMM has powered on the storage node. The CMM event log might also indicate why the node canister has no power. Otherwise try re-seating the canister. If the state persists, follow the hardware replacement procedures for the parts in the following order: node canister, control enclosure. ► **ON SOLID** – The canister is powered on. This is the normal state. ► **BLINK** – The canister is in a standby state after being powered down. Use the CMM to power on the canister. **Note:** The Power LED can also be in the BLINK state after the canister has been inserted and has been discovered by the CMM but before the V7000 code has booted. ► **FAST BLINK** – The node canister is in the process of communicating with the CMM during initial insertion of the canister. If the canister remains in this state for more than 10 minutes, Try re-seating the canister. If the state persists, follow the hardware replacement procedure for the node canister

2. If the power LED is on solid, consider the state of the canister status LED **9** for each node canister**,** as shown in Table 13-3 on page 527.

Table 13-3 Node canister status LED statues

LED layout	LED Name	Color	States
9	 Canister status	Green	► **OFF** – The IBM Flex System V7000 system code is not running. If the Power LED is ON, the operating system has not started the IBM Flex System V7000 system code yet. This state is normal just after the canister is powered on. If it persists for more than 10 minutes, check the canister status using the CMM. ► **ON SOLID** – The canister is active and able to perform I/O as part of a IBM Flex System V7000. The IBM Flex System V7000 management GUI is available if any node canister in the system has its status LED ON. Do not power off, or remove, a node canister whose status LED is on solid. You might lose access to data or corrupt volume data. Follow the proper procedure to shut down a node so that access to data is not compromised. ► **BLINK** – The canister is in candidate or service state and is not participating in I/O. If possible, use the recommended action in the event log event log shown by the management GUI to service the node canister. If this is not possible, use the service assistant to determine the state of the node canister and resolve any critical node errors that stop the canister from becoming active.

3. If the power LED is on solid, consider the states of the fault LEDs 3 and 4 for each node canister, as shown in Table 13-4

Table 13-4 Node canister fault LED states

LED layout	LED Name	Color	States
3	 Canister fault	Amber	► **OFF** – The node canister is operating correctly, or it has not been possible to isolate an issue to the node canister. ► **ON SOLID** – The node canister has failed and should be replaced. If it is available, use the IBM Flex System V7000 management GUI event log fix procedures to guide the replacement.
4	 Internal fault	Amber	There is a fault on one or more of the replaceable parts in the canister, it has been determined that the part needs replacing. **Note:** This LED only applies to the network adapters. (There may be one or two of them.) Review the IBM Flex System V7000 Storage Node event logs or the CMM to identify the network adapter that failed. ► **OFF** – There are no failures isolated to internal components of the canister ► **ON SOLID** – Replace the failing network adapter

4. If the canister fault LED and the internal fault LED do not indicate a problem, consider the state of the battery LEDs **5, 6**, and **7**, as shown in Table 13-5 on page 528 on each node canister. The canister will not run as part of the IBM Flex System V7000 unless it is sure it can save its state and cache data in the event of power loss to the canister. See Understanding the battery operation for the node canister.

Table 13-5 Node canister battery LED states

LED layout	LED Name	Color	States
5	Battery Status	Green	► **OFF** – Indicates the battery is not in a state where it can support a save of cache and system state data. This is an error condition. Refer to the battery fault LED and the node status provided by the service assistant for further information. ► **ON SOLID** – Indicates the battery is fully charged and can support a save of cache and system state data. This is the normal state. ► **BLINK** – Indicates the battery is charging and can support at least one save of cache and system state data. No action is required; the node canister can function as an active member of the system in this state. ► **FAST BLINK** – Indicates the battery is charging, but cannot yet support a save of cache and system state data. The node canister cannot operate as an active member of the system in this state. No action is required; when there is sufficient charge to save the cache and system state data, the node canister automatically becomes active in the system.
6	Battery fault	Amber	► **OFF** – No faults have detected with the battery ► **ON SOLID** – A fault has been detected with the battery. The battery needs to be replaced. ► **Attention:** If your system is powered on and performing I/O operations, go to the management GUI and follow the fix procedures. Removing a node without the assistance of the fix procedures can result in loss of data or access to data. ► **BLINK** – The battery has been identified by user action.
7	Battery in use	Green	► **OFF** – The battery is not providing power to the node canister. This is the normal state ► **FAST BLINK** – The system is currently in the process of saving cache and system state data to its internal storage device using battery power. Do not remove, or perform any other service action on the canister until the LED stops blinking. If the LED continues to fast blink for more than fifteen minutes, perform the following actions: – Re-seat the node canister. – Re-seat the battery in the node canister. – Replace the battery. – Replace the node canister.

5. Table 13-6 describes the SAS port status and activity LEDs, canister activity LED, and identify LEDs on the canisters.

Table 13-6 Node canister SAS port and control enclosure LED states

LED layout	LED Name	Color	States
1	SAS port status	Amber	► **OFF** – There are no reported faults on the SAS link. This is the normal state. ► **ON SOLID** – There is no connectivity on the link ► **SLOW BLINK** – The port has been disabled and will not service SAS traffic ► **BLINK** – One or more of the narrow ports of the wide SAS port link have failed, the port is not operating at full capacity. If the LED is not OFF, use the IBM Flex System V7000 management GUI, if available, and use event log fix procedures to troubleshoot. These procedures verify that the SAS cable is seated correctly at each end, it is working, or if one of the canisters it connects to needs to be replaced.

LED layout	LED Name	Color	States
2	SASport activity	Green	► **OFF** – Power is not present or there is no SAS link connectivity established. Check that the SAS cable is correctly connected at each end, and that the canisters it connects to are both powered on. The cable may need replacing. ► **ON SOLID** – There is at least 1 active SAS link in the wide port established and there is no external port activity. This is a normal state and no action is required. ► **FLASHING** – The expansion port activity LED should flash at a rate proportional to the level of SAS port interface activity as determined by the canister. The port also blinks when routing updates or configuration changes are being performed on the port. This is a normal state and no action is required.
10	Canister activity	Green	► **OFF** – There is no host I/O activity. ► **FLASHING** – The canister is actively processing input/output traffic (host I/O activity).
11	t Enclosure Fault	Amber	► **OFF** – There are no isolated failures on the storage enclosure. ► **ON SOLID** – There are one or more isolated failures in the storage enclosure requiring service or replacement. See the IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide PDF on the CD for troubleshooting procedures.
12	Check log	Amber	► **OFF** – There are no conditions that require the user to login to the management interface and review the error logs. ► **ON SOLID**– The system requires the attention of the user via one of the management interfaces. There are multiple reasons that the Check Log LED could be illuminated.
13	Canister or control enclosure identify	Blue	► **OFF** – There has been no request to identify the canister or enclosure. ► **ON SOLID** – The canister or enclosure has been identified in request to user action. ► **BLINK** – This occurs during power on and power-on self-test (POST) activities and when a command is being run from a USB flash drive

13.2.2 Understanding the system state using the expansion enclosure LEDs

The LEDs provide a general idea of the expansion canister status. You can obtain more detail from the management GUI and the Service Assistant. Examine the LEDs when you are not able to access the management GUI or the service assistant, or when the system is not showing any information about a device.

The CMM also shows status information about your expansion canisters, which you can use when the management GUI or the service assistant cannot be accessed. The CMM indicates the current status of many of the LEDs on your IBM Flex System V7000.

The procedure shows the status for the expansion canisters. It does not show the status for the drives.

IBM Flex System V7000 expansion enclosure includes two expansion canisters that are located in the rear of the expansion enclosure. The expansion canister has no controls and does not include battery unit. Figure 13-3 and Table 13-7 show expansion canister LEDs you can use to determine the system state and appropriate actions to take to resolve the issues.

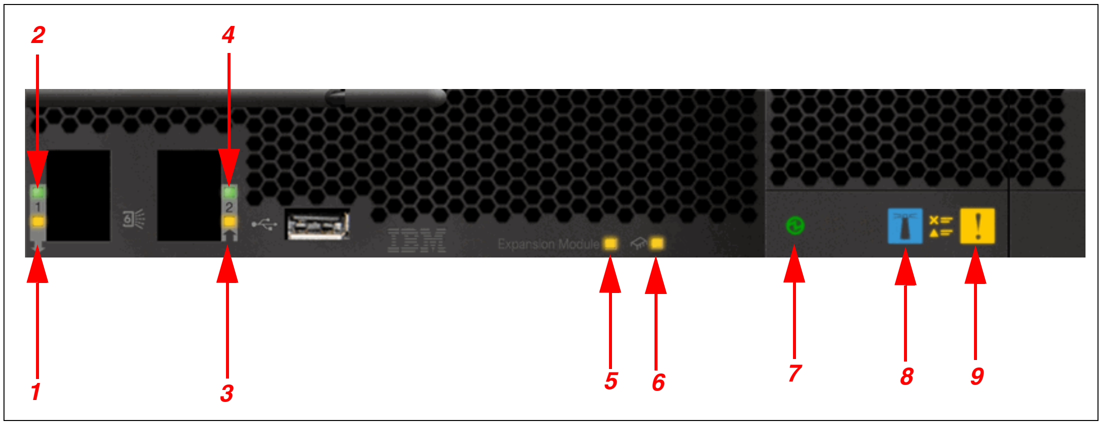


Figure 13-3 Expansion enclosure LEDs layout

Table 13-7 Expansion canister LEDs

LED layout	LED name	Color	Color
1 3	!	SAS port status	Amber
2 4	N/A	SAS port activity	Green
5	!	Expansion canister fault	Amber
6		Expansion canister internal fault	Amber
7		Power	Green

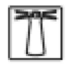
LED layout	LED name	Color	Color
8		Canister or expansion enclosure identify	Blue
9		Expansion enclosure fault	Amber

Table 13-8 defines the various LED states for the canisters in the IBM Flex System V7000 expansion canister.

Table 13-8 Definition of a node canister LED states

State	Description
Off	The LED is continuously not lit or off
Blinking slowly	The LED turns on and off at a frequency of 1 Hz: It is on for 500 ms, then off for 500 ms, then repeats.
Blinking	The LED turns on and off at a frequency of 2 Hz: It is on for 250 ms, then off for 250 ms, then repeats.
Blinking fast	The LED turns on and off at a frequency of 4 Hz: It is on for 125 ms, then off for 125 ms, then repeats.
On	The LED is continuously lit or on
Flashing	The LED is lit to indicate some activity, then turns off. The rate and duration that the LED is lit depends on the rate and duration of the activity.

1. Start by considering the states of the power LED **7** for each 4939 expansion canister, as shown in Figure 13-3 on page 530 and Table 13-9 on page 532. Continue to troubleshoot the canister until the power LED is ON.

Table 13-9 Expansion canister Power LED states

LED layout	LED Name	Color	States
7	Power	Green	► **OFF** – There is no power to the expansion canister. Make sure that the CMM has powered on the expansion canister. The CMM event log might indicate why the expansion canister has no power. Try reseating it. If the state persists, follow the hardware replacement procedures for the parts in the following order: – Expansion canister. – Expansion enclosure. ► **ON SOLID** – The expansion canister is powered on. This is the normal state. ► **BLINK** – The expansion canister is in a powered down state. Use the CMM to power on the expansion canister. ► **FAST BLINK** – The management controller is in the process of communicating with the Chassis Management Module (CMM) during initial insertion of the expansion canister. If the expansion canister remains in this state for more than 10 minutes, try reseating it. If the state persists, follow the hardware replacement procedure for the expansion canister.

2. If the power LED is on, consider the states of the status and fault LEDs for each expansion canister, as shown in Table 13-10.

> **Note:** Although it is useful to understand the meaning of the expansion canister LEDs, service actions should always start by examining the IBM Flex System V7000 management GUI event log and running the Fix Procedures for the reported errors in the order suggested

Table 13-10 Expansion canister status LED states

LED layout	LED Name	Color	States
5	Expansion canister fault	Amber	► **OFF** – The expansion canister is operating correctly, or it has not been possible to isolate the issue to make it the expansion canister. ► **ON SOLID** – The expansion canister has failed and should be replaced. If it is available, use the IBM Flex System V7000 management GUI event log fix procedures to guide the replacement.
8	Expansion canister identify	Blue	► **OFF** – The expansion canister is not in identify state by the controller management system ► **ON SOLID** – The expansion canister has been identified in response to the controller management system ► **BLINK** – This occurs during power on and power-on self-test (POST) activitie
9	Expansion enclosure fault	Amber	► This LED is not used and it is always off.

LED layout	LED Name	Color	States
1 3	SAS port status	Amber	► **OFF** – There have been no faults or conditions detected by the expansion canister on the SAS port or down stream device connected to the port ► **ON SOLID** – There has been a fault condition isolated by the expansion canister on the external SAS port ► **SLOW BLINK** – The port has been disabled and will not service SAS traffic ► **BLINK** – One or more of the narrow ports of the SAS links on the wide SAS port link have failed, the port is not operating as a full wide port
2 4	SAS port activity	green	► **OFF** – Power is not present or there is no SAS link connectivity established ► **ON SOLID** – There is at least 1 active SAS link in the wide port established and there is no external port activity ► **FLASHING** – The expansion port activity LED should flash at a rate proportional to the level of SAS port interface activity as determined by the expansion canister. The port shall also blink when routing updates or configuration changes are being performed on the port

13.2.3 Power-on self-test

When you turn on the system, the node canisters perform self-tests. A series of tests is performed to check the operation of components and some of the options that have been installed when the units are first turned on. This series of tests is called the power-on self-test (POST).

If a critical failure is detected during the POST, the software is not loaded and the fault LED is illuminated. To determine if there is a POST error on a canister, go to 13.2.1, “Understanding the system state using the control enclosure LEDs” on page 524 and 13.2.2, “Understanding the system state using the expansion enclosure LEDs” on page 529

When the software is loaded, additional testing takes place, which ensures that all of the required hardware and software components are installed and functioning correctly.

13.2.4 Powering on using LED indicators

After installing the enclosures in the IBM Flex System Enterprise Chassis, use the following procedure to verify the enclosure power on status using the light emitting diodes (LEDs).

Before you begin

- Make sure that the Flex System chassis is powered on. If the IBM Flex System Enterprise Chassis is not powered on, the enclosures are not able to power on.
- Observe the power on LED state for each canister in each enclosure. See Table 13-2 on page 526 and Table 13-9 on page 532 for the power on LED states. The power on LED for each canister must be in the blink state before the enclosures will power on.

You might have to wait up to ten minutes after installing the enclosure in the chassis before the power on LEDs reach the blink state and can be automatically powered on. If the enclosure does not reach the blink state after waiting ten minutes, see IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide PDF on the CD

Attention: Do not operate the system when the drive assemblies are missing. Drive assemblies that are missing disrupt the airflow; the drives do not receive sufficient cooling. You must insert blank carriers into any unused drive bays.

Powering on a control enclosure procedure

1. Powering on a control enclosure
2. When the power on LEDs reach the blink state, the control enclosure is powered on automatically. The green power on LEDs are on solid when the enclosure is powered on successfully.
3. Verify that no faults are detected on the control enclosure after powering on. Use Table 13-2 on page 526 to verify the state of the LEDs on the control enclosure. See the IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide PDF on the CD if problems are encountered. Figure 13-2 on page 526 shows the location of the LEDs on the front of the control enclosure.

Attention: Do not continue if any faults are indicated by the LEDs.

Powering on an expansion enclosure

1. When the expansion enclosure power on LEDs reach the blink state, the expansion enclosure is powered on automatically. The green power on LEDs are on solid when the enclosure is powered on successfully. See Table 13-9 on page 532.
2. Use the information in Table 13-10 on page 532 to verify the state of the light emitting diodes (LEDs) on the expansion enclosure. Verify that no faults are detected. See the IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide PDF on the CD if problems are encountered.

Attention: Do not continue if any faults are indicated by the LEDs.

13.3 Monitoring system status and health

In this section we describe the procedure to monitor the system status and the health.

13.3.1 Using FSM for Status and Health

IBM Flex System Manager provides a set of tools that you can use to monitor and manage the status and health of resources in your environment from a single interface.

Note: The Chassis Manager view in the management software web interface is the primary interface for selecting managed resources and performing management tasks on those resources. However, some of the procedures in this section include steps that instruct you to use the navigation area, which is hidden by default in the web interface

The navigation area provides links to tasks and task categories such as Resource Explorer, Availability, Inventory, System status and Health Summary. To open the navigation area, click the tab with the arrow icon on the left side of the screen, as shown in Figure 13-4 on page 535.

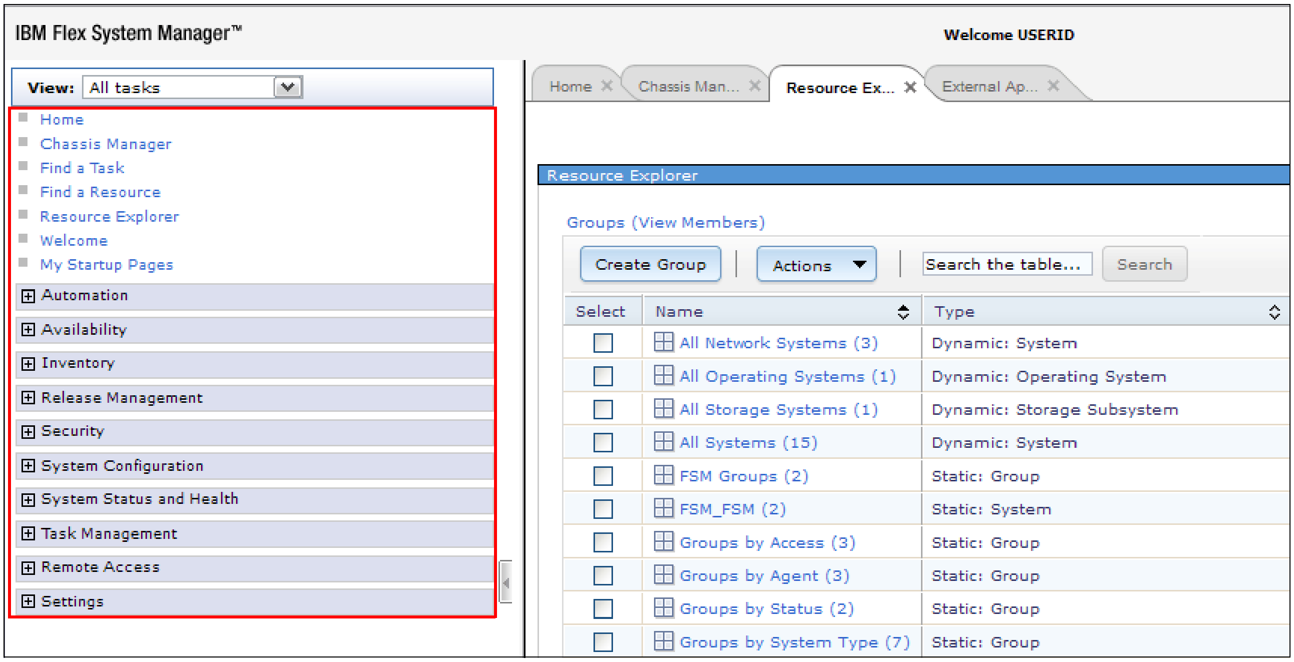


Figure 13-4 FSM - tasks management and navigation

IBM Flex System Manager automatically retrieves and displays the status of systems that have been discovered. You can display this information using one of the System Status and Health tasks, by navigating to a specific resource in IBM Flex System Manager, or by using the command-line interface. After the first discovery of the resources, IBM Flex System Manager automatically monitors the system and its resources, as shown in Figure 13-5.

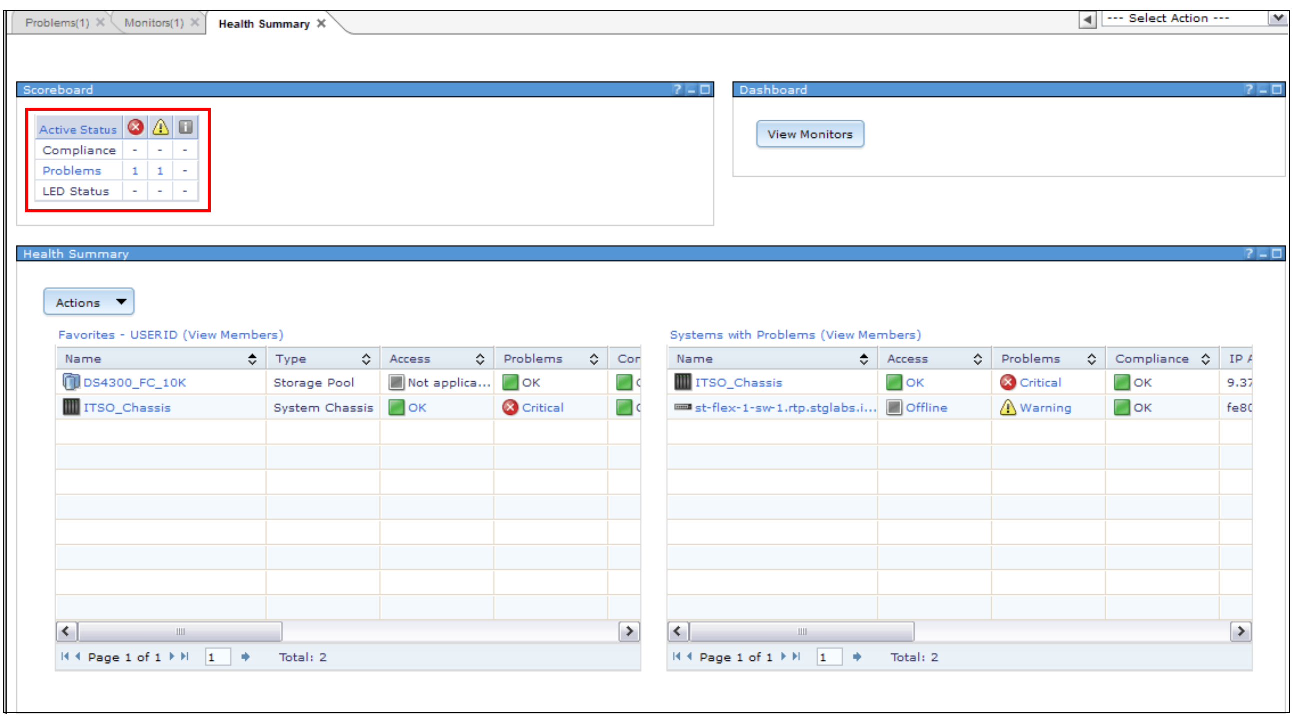


Figure 13-5 FSM - System Status and Health Scoreboard

The System Status and Health tasks simplify your work as a system administrator by providing consolidated views of the health and status of your entire environment. Specifically, the Health Summary task displays information about the resources that IBM Flex System Manager has discovered. You can customize the content that is displayed on the System Status and Health pages with information that is specific to your business processes and organizational structure. In addition, you can view detailed real-time data by adding monitors to the dashboard.

Use the following System Status and Health tasks to view the system's status, as shown in Figure 13-6

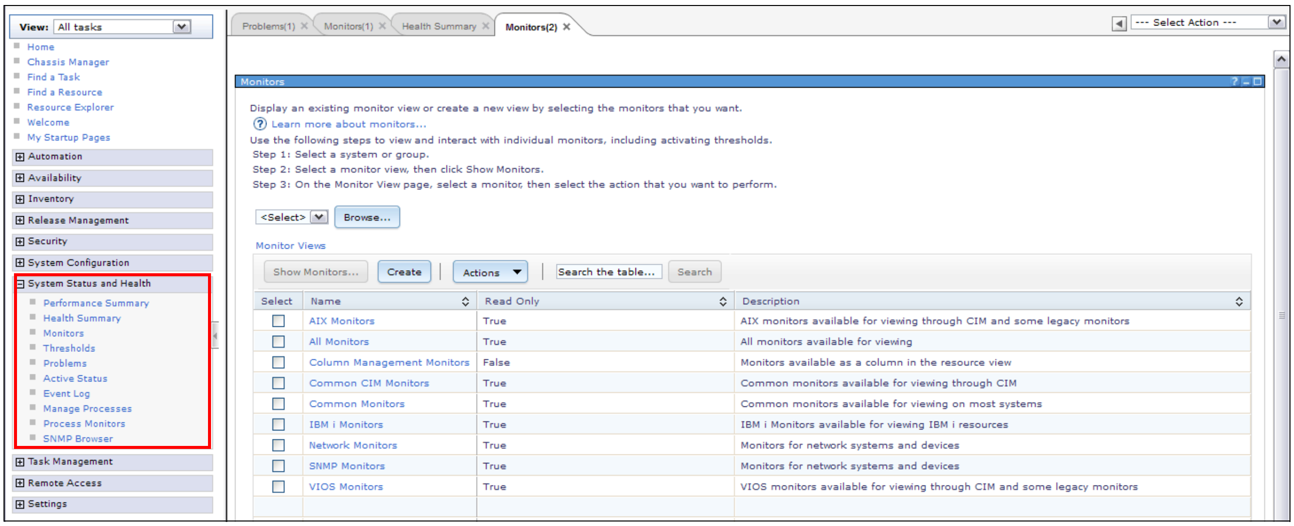


Figure 13-6 FSM - System Status and Health tasks

- Performance Summary

  The Performance Summary task displays available monitors with which you can examine the servers, virtual servers, or operating systems that you specify. Use this page to dynamically view, monitor, and work with processor pools, memory pools, networks, and other critical areas of your environment.

- Health Summary

  The Health Summary task displays several resource-monitoring tools on a single page. Together, these tools provide a single, consolidated interface with which you can quickly view the status of important areas of your environment, monitor critical resources, and view the contents of user-defined health summary groups.

- Monitors

  The Monitors task provides the tools that you need to retrieve real-time status and quantitative data for specific properties and attributes of resources in your environment. You can also set thresholds for the monitors, graph the data that monitors retrieve, and drill down to quickly view the status of resources for each system and the name of the monitor so that you can view its properties.

- Thresholds

  The Thresholds task offers a consolidated view of all the thresholds that you have created to monitor the dynamic properties of your resource. This task saves you from searching for them all in the Monitors task.

- Problems

The Problems task lists only those elevated status set entries that are problem status set entries that have been reported to IBM Flex System Manager. The data listed here is a subset of what is listed on the Active Status page. From the Problems page, you can navigate to, display, and manage these status set entries. You can also ignore entries to prevent IBM Flex System Manager from collecting them or delete the entries to remove them from the list until they occur again.

- Active Status

  The Active Status task lists all of the status set entries, including problems and compliance issues, that have been reported to IBM Flex System Manager. From the Active Status page, you can navigate to, display, and manage these status set entries. You can also ignore entries to prevent IBM Flex System Manager from collecting them or delete the entries to remove them from the list until they occur again.

- Event Log

  An event is an occurrence of significance to a task or resource. Examples of events include operation completion, hardware component failure, or a processor threshold being exceeded. The Event Log task displays all events that the management server receives from any resource for which you can view events.

- SNMP Browser

  The SNMP Browser task enables you to display, inspect, and perform certain actions on SNMP-compliant resources, including selecting the targets on which an SNMP job will run.

For implementing details, go to *Implementing Systems Management of IBM PureFlex System*, SG24-8060

13.4 Managing storage nodes

In addition to the IBM Flex System V7000 Storage Node management GUI, you can use the IBM Flex System Manager management software to perform management tasks on a Flex System V7000.

The Flex System V7000 provides virtualized storage within an IBM Flex System environment, and provides shared storage capacity to compute nodes by virtualizing internal disk drives and external Fibre Channel storage system capacity.

After a new IBM Flex System V7000 Storage Node is installed in an IBM Flex System Enterprise Chassis, the Chassis Map in the management software web interface shows the installed Flex System V7000 and provides a link to the Launch IBM Flex System V7000 setup wizard task in the menu that enables you to set up the storage node. See Using the Flex System Manager (FSM) to setup your system 4.2.1, “Using FSM for initial setup” on page 154.

After you complete the Setup wizard, the management software discovers the Flex System V7000 and collects inventory automatically. After the Flex System V7000 is managed by the management software, you can use the management software to perform management tasks, including monitoring, configuring, virtualizing, and troubleshooting the storage node. You can use the IBM Flex System Manager management software web interface to manage and monitor storage performance.

13.4.1 Using FSM Chassis Manager page to manage storage node

The Chassis Manager page and Chassis Map view provide hardware and table views of all discovered resources in a managed chassis. Use the Chassis Manager view in the management software web interface to view Flex System V7000 details and perform management tasks.

> **Note:** If an IBM Flex System V7000 Storage Node is not discovered and inventoried automatically, use the IBM Flex System Manager management software to discover, inventory, and manage the Flex System V7000 manually. For more information, refer to : `http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_product_page.html`

To access a Flex System V7000 from the Chassis Manager page, complete the following steps:

1. From the Home page, click the Chassis Manager tab.
2. From the Chassis Manager page, click the chassis that contains the Flex System V7000 storage node. The Chassis Map view opens.
3. From the Chassis Map, click the Flex System V7000 storage node to select it. The details for the storage node are displayed.
4. Click the Actions menu for a list of storage viewing and management options, as shown in Figure 13-7 on page 539

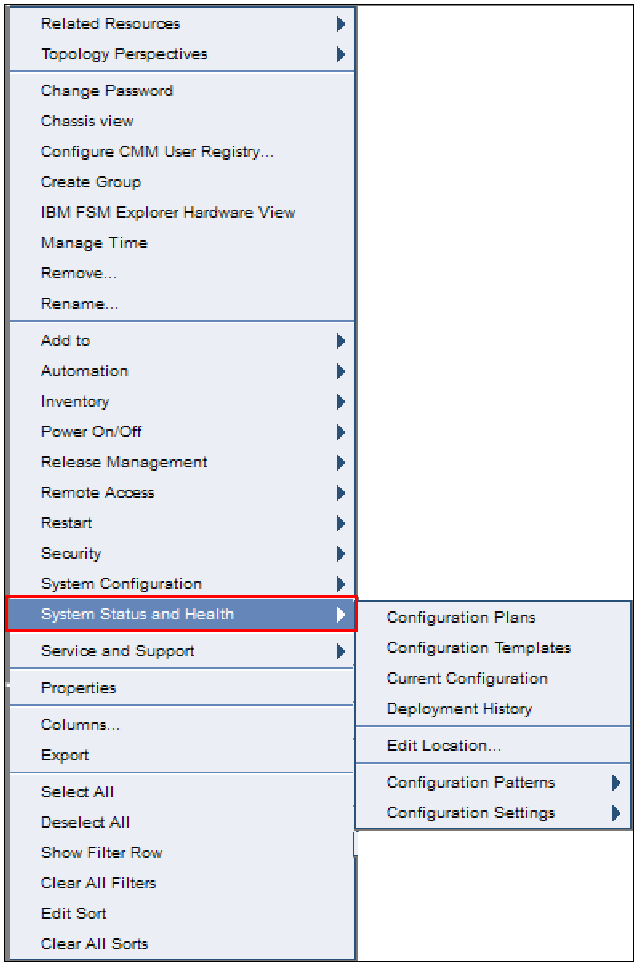


Figure 13-7 FSM - access a Flex System V7000 from the Chassis Manager page

> **Note:** You can launch the Flex System V7000 management GUI from the Chassis Map view. After you select the storage node in the Chassis Map view, click Launch IBM Flex System V7000

13.4.2 Using FSM Storage Management to manage storage node

The Storage Management summary page in the management software web interface enables you to view storage capacity details and provides links to storage monitoring tasks.

To access the Storage Management summary page, from the Home page, click the Plug-ins tab and click Storage Management. The Storage Management summary page opens, as shown in Figure 13-8 on page 540

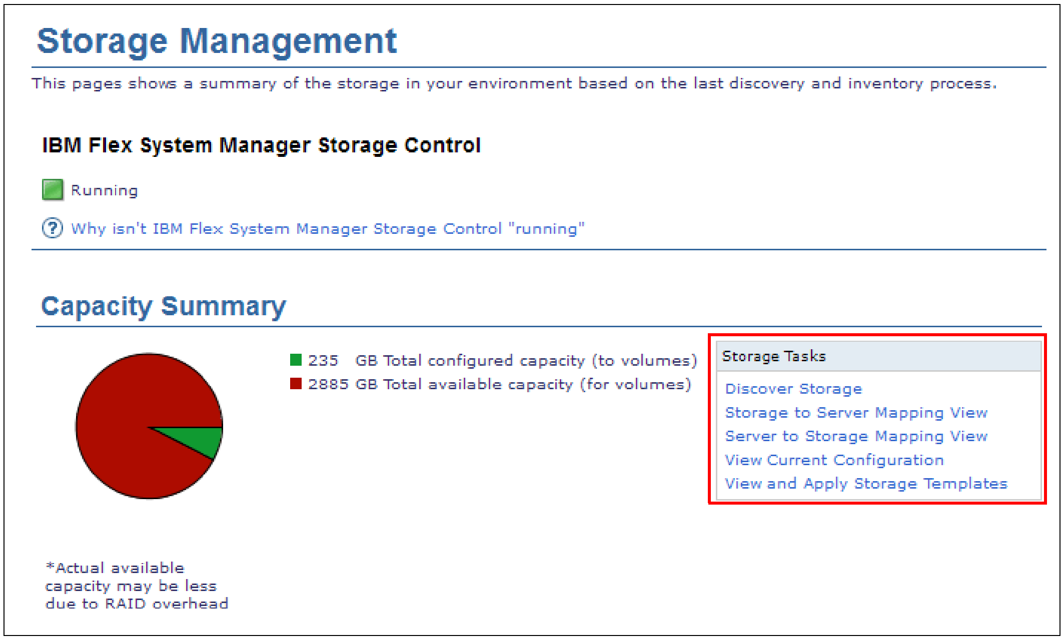


Figure 13-8 FSM - Storage Management

The following links are available under Storage Tasks:

- Discover Storage
- Storage to Server Mapping View
- Server to Storage Mapping View
- View Current Configuration
- View and Apply Storage Templates

For implementation details, refer to 3.3, “Flex System Manager (FSM)” on page 110.

13.5 Configuration backup and restore process

Configuration backup is the process of extracting configuration settings from a clustered system and writing it to disk. The configuration restore process uses backup configuration data files for the system to restore a specific system configuration. Restoring the system configuration is an important part of a complete backup and disaster recovery solution.

Attention: Only the data that describes the system configuration is backed up. You must back up your application data using the appropriate backup methods.

For complete disaster recovery, regularly back up the business data that is stored on volumes at the application server level or the host level.

To enable routine maintenance of the IBM Flex System V7000 Storage Node clustered systems, the configuration settings for each system are stored on each node. If power fails on a system or if a node in a system is replaced, the system configuration settings are automatically restored when the repaired node is added to the system. To restore the system configuration in case of a disaster (if all nodes in a system are lost simultaneously), plan to back up the system configuration settings to tertiary storage. You can use the configuration backup functions to back up the system configuration.

To generate a configuration backup using the CLI, go to:
http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/svc_clustconfbackuptsk_1e4k69.html

Table 13-11 describes the three files that are created by the backup process.

Table 13-11 File names created by the backup process

File name	Description
svc.config.backup.xml	This file contains your cluster configuration data.
svc.config.backup.sh	This file contains the names of the commands that were issued to create the backup of the cluster.
svc.config.backup.log	This file contains details about the backup, including any error information that might have been reported.

13.6 Software upgrade

The system upgrade process involves the upgrading of your entire Flex System V7000 Storage Node environment. The amount of time can vary depending on the amount of preparation work required and the size of the environment. Generally allow slightly over 1 hour for an upgrade.

13.6.1 Upgrading the System Software

Before you upgrade, become familiar with the prerequisites and tasks involved. Decide whether you want to upgrade automatically or upgrade manually. During an automatic upgrade procedure, the clustered system upgrades each of the nodes systematically. The automatic method is the preferred procedure for upgrading software on nodes. However, you can also upgrade each node manually.

Attention:

- If you experience failover issues with multipathing driver support, resolve these issues before you start normal operations.
- Ensure that you have no unfixed errors in the log and that the system date and time are correctly set. Start the fix procedures, and ensure that you fix any outstanding errors before you attempt to concurrently upgrade the code.
- If you want to upgrade without host I/O, shut down all hosts before you start the upgrade.
- Some code levels support upgrades only from specific previous levels, or the code can be installed only on certain hardware types. If you upgrade to more than one level above your current level, you might be required to install an intermediate level. For example, if you are upgrading from level 1 to level 3, you might need to install level 2 before you can install level 3. For information about the prerequisites for each code level, see the IBM Support website.

Upgrading the software automatically

During the automatic upgrade process, each node in a system is upgraded one at a time, and the new code is staged on the nodes. While each node restarts, there might be some degradation in the maximum I/O rate that can be sustained by the system. After all the nodes

in the system are successfully restarted with the new code level, the new level is automatically committed.

During an automatic code upgrade, each node of a working pair is upgraded sequentially. The node that is being upgraded is temporarily unavailable and all I/O operations to that node fails. As a result, the I/O error counts increase and the failed I/O operations are directed to the partner node of the working pair. Applications do not see any I/O failures. When new nodes are added to the system, the upgrade package is automatically downloaded to the new nodes from the IBM Flex System V7000 Storage Node.

The upgrade can normally be performed concurrently with normal user I/O operations. However, there is a possibility that performance could be impacted. If any restrictions apply to the operations that can be performed during the upgrade, these restrictions are documented on the Flex System V7000 Storage Node website that you use to download the upgrade packages. During the upgrade procedure, the majority of configuration commands are not available.

This automatic procedure provides a unified mechanism to upgrade the entire system in a coordinated process with no user intervention. Excluding drives, when a node is rebooted as part of the clustered system upgrade, the system checks that it is at the correct level. When the system detects that the hardware is not running at the expected level, the system applies the correct firmware.

Note: Use the Software Upgrade Test Utility to help you find potential upgrade-related problems. You can download the most current version of this tool at the following website:

When you are ready to upgrade, click **Settings** → **General** → **Upgrade Software** in the management GUI and follow the instructions.to complete, as shown in Figure 13-9. See “Upgrade Procedure” on page 543.

Figure 13-9 Upgrading software in the management GUI

Monitor the upgrade information in the management GUI to determine when the upgrade is complete.To determine when your upgrade process has completed, you are notified through the management GUI. If you are using the command-line interface, issue the `lssoftwareupgradestatus` command to display the status of the upgrade.

Upgrading the software manually

During an automatic upgrade procedure, the Flex System V7000 Storage Node clustered system upgrades each of the nodes systematically. The automatic method is the preferred procedure for upgrading software on nodes; however, to provide more flexibility in the upgrade process, you can also upgrade each node manually.

Note: Immediately prior to a manual upgrade, the cluster must be prepared by issuing the **`svctask applysoftware -prepare -file svc_software_package`** command in the CLI. Refer to the below link for more information.

`http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/svc_userpacedsftwrupgrade_419938.html`

During this manual procedure, the upgrade is prepared, you remove a node from the system, upgrade the code on the node, and return the node to the system. You repeat this process for the remaining nodes until the last node is removed from the system. Every node must be upgraded to the same code level. You cannot interrupt the upgrade and switch to installing a different level. When the last node is returned to the system, the system completes the upgrade and starts running the new level of code.

Concurrent code update of all components is supported through the standard Ethernet management interfaces.

Upgrade Procedure

To upgrade the IBM Flex System V7000 Software, complete the following steps:

1. Open a supported web browser and point your web browser at the management IP address that was set during the initial setup process, refer to Chapter 4, "IBM Flex System V7000 Storage Node Initial Configuration" on page 149.
2. Log in to your IBM Flex System V7000 Storage Node management GUI(Disperser and default password: passw0rd) From IBM Flex System V7000 management GUI home page, go to **Settings** → **General**, as show in Figure 13-10.

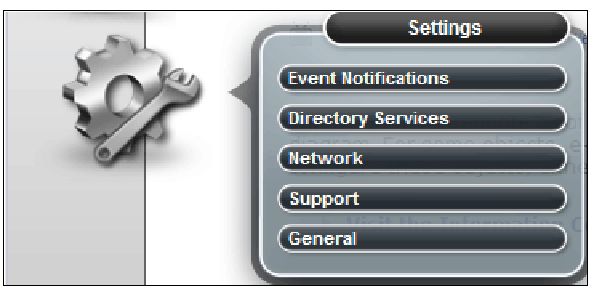


Figure 13-10 Management GUI - settings

3. The current software level is displayed on the top of the and you may want to check for new updates before you start to upgrade, go to **Upgrade Software** → **Check for Updates** to contact upgrade server, as shown in Figure 13-11 on page 544.

   Note: If the control enclosure doesn't have internet access, your Check for Updates will fail with the following message: Unable to contact upgrade server. If your web browser is behind a firewall or does not have Internet access, download the upgrade package and manually install it on the system.

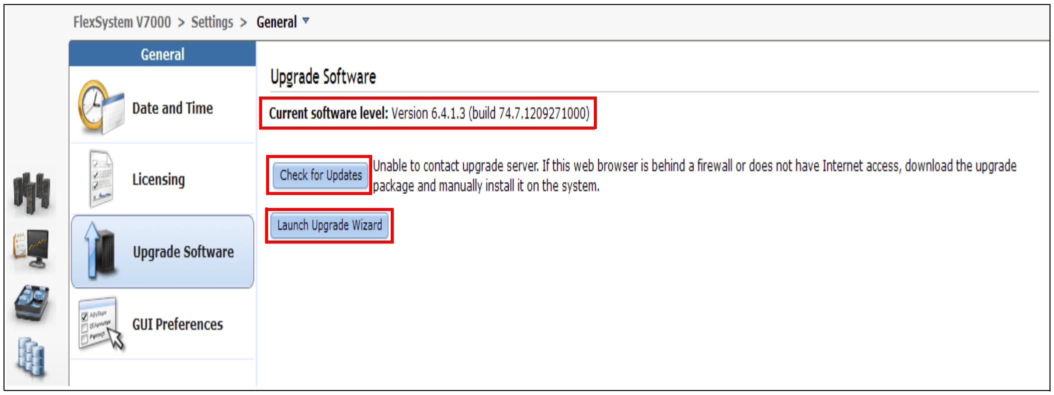


Figure 13-11 Check for software updates

4. Click **Launch Upgrade Wizard** as shown in Figure 13-11 to download the upgrade test utility. This utility checks for known issues which might prevent the upgrade from completing successfully. Click Download link to download the test utility directly from the web server. If your web browser is behind a firewall or does not have Internet access, download the package manually to your laptop, as shown in Figure 13-12.

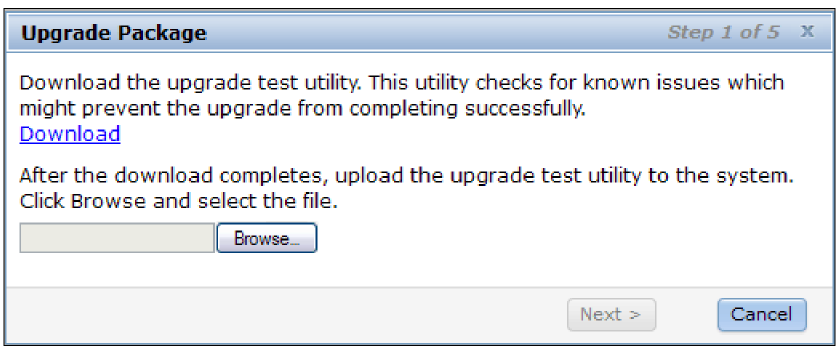


Figure 13-12 Step 1 of 5 - Download upgrade test utility from the download web site

5. After the download completes, click **Next** to upload the upgrade test utility to the system, as shown in Figure 13-13.

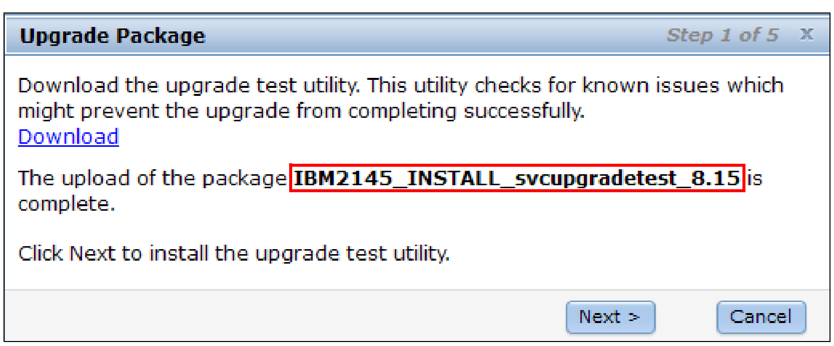


Figure 13-13 Step 1 of 5 - Upload upgrade test utility to the system

6. Enter the correct version info that you want to upgrade to. In this example, our version is 6.4.1.3. The upgrade test utility checks the system to ensure it is ready for an upgrade to this version. Click **Next** to run the upgrade test utility, as shown in Figure 13-14 on page 545

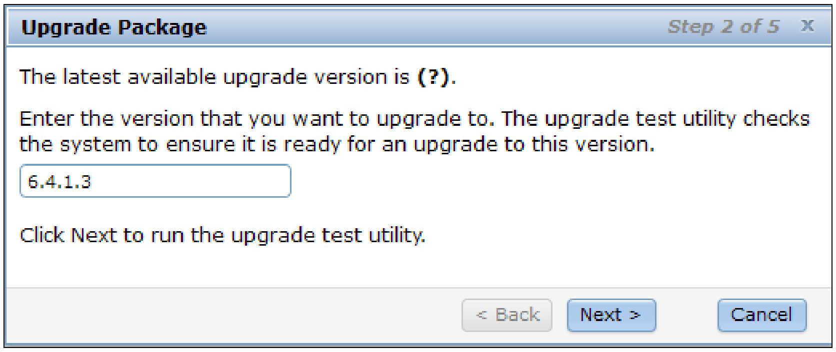


Figure 13-14 Step 2 of 5 - Run upgrade test utility

> **Note:** The test may take several minutes to complete. If the upgrade test utility used didn't match with the upgrade version you want to upgrade to, you will receive a message similar to the following message
>
> This version of svcupgradetest cannot test the currently running version of cluster code. This cluster has software version 6.4.1.3 installed. Please check the support website to see if a newer version of this tool is available.
>
> It's highly recommended to run upgrade test utility before you perform the upgrade software task. Obtain the appropriate version of the upgrade test utility and try it again before continuing the software upgrade.

7. Click **Next** to check for any dependent volumes and to proceed with the upgrade.
8. Click the Download link to download the latest upgrade package from the web server. If your web browser is behind a firewall or does not have Internet access, download the package manually to your laptop.
9. Click **Next** to upload the upgrade package to the system and next click Browse to select a file, as shown in Figure 13-15.

> **Note:** you can download the IIBM Flex System V7000 software upgrade package, or you can browse and upload the software upgrade package from the location where you saved it. Automatic upgrade process provides a unified mechanism to upgrade the entire system in a coordinated process with no user intervention.

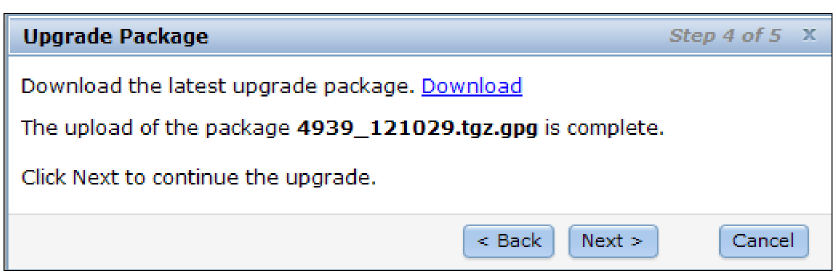


Figure 13-15 Step 4 of 5 - Upload the package to the system

10.I Click Next to continue the upgrade. Select Automatic upgrade and click OK, as shown in Figure 13-16 on page 546.

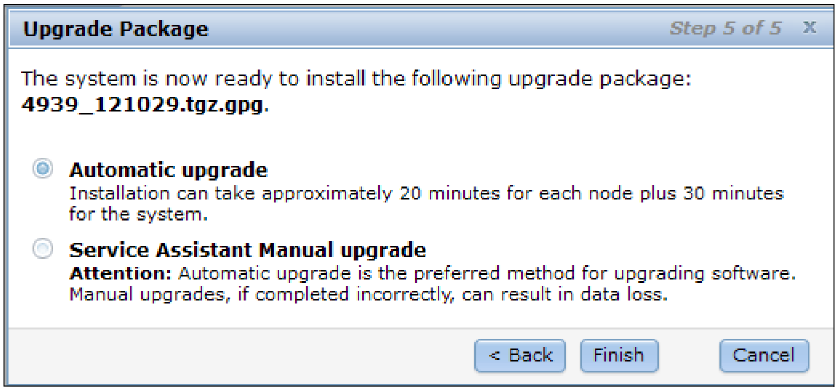


Figure 13-16 Step 5 of 5 Start the upgrade h

11. Click Finish to complete your software upgrade. You receive message that inform you that first node, then the other, has been upgraded. When both nodes have been rebooted, you have completed your IBM Flex System V7000 Storage Node software upgrade, as shown in Figure 13-17.

> **Note:** System Health Status changes from green to yellow as soon as the upgrade started on the first node canister, as shown in Figure 13-17. It will change from yellow to green again when the upgrade on the 2nd node is complete, as shown in Figure 13-18 on page 547

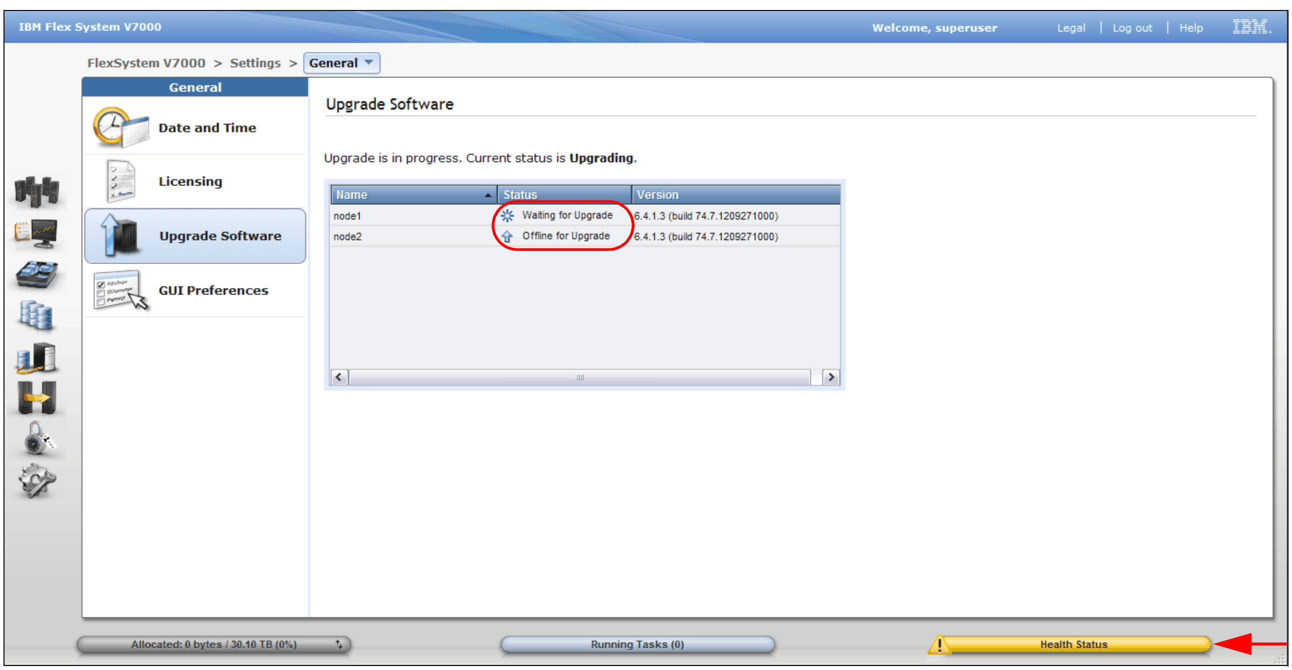


Figure 13-17 Monitoring upgrade status

After the upgrade on the 2nd node is complete, the health status will turn green as shown in Figure 13-18 on page 547.

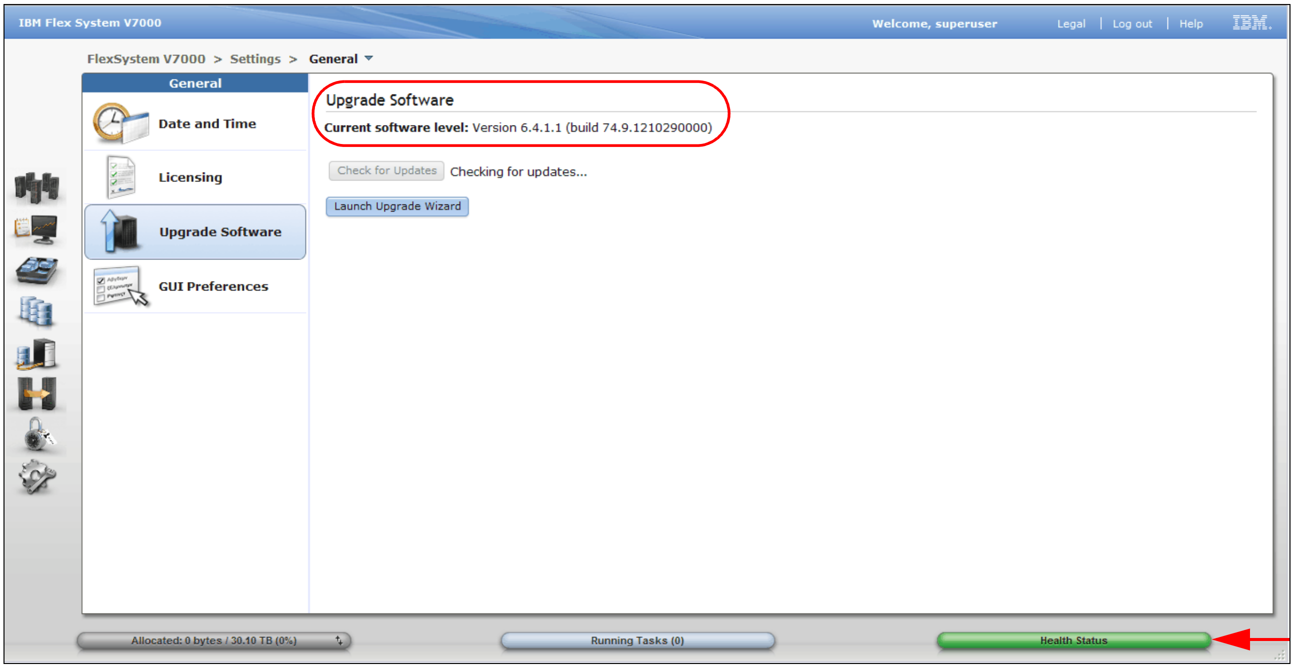


Figure 13-18 Software upgrade complete

13.7 Drive Firmware upgrade

You can upgrade a drive by downloading and applying firmware updates by using the command-line interface (CLI). The time required to perform drive upgrade varies depending on the number of drives to be upgraded and the capacity of these drives. You need to plan your drive firmware upgrade task accordingly.

This procedure upgrades firmware on a drive that is in the control enclosure or in the expansion enclosures. If the upgrade would cause any volumes to go offline, the force option is required.

To upgrade drive firmware, follow the following procedure:

1. Run the following command for the drive that you are upgrading

   Command: lsdependentvdisks -drive drive_id

   If any volumes are returned, continuing with this procedure takes the volumes offline. To avoid losing access to data, resolve any redundancy errors to remove this problem before you continue with the upgrade procedure.
2. Locate the firmware upgrade file at the following website:

   www.ibm.com/storage/support/flexsystemFlex System /v7000<< Author comment - need to verify if this web download link is valid>>

   This website also provides a link to the Software Upgrade Test Utility. This utility indicates if any of your drives are not running at the latest level of firmware.
3. Using scp or pscp, copy the firmware upgrade file and the Software Upgrade Test Utility package to the /home/admin/upgrade directory by using the management IP address.

4. Run the `applydrivesoftware` command. You must specify the firmware upgrade file, the firmware type, and the drive ID.

Command: `applydrivesoftware -file name -type firmware -drive drive_id`

To apply the upgrade even if it causes one or more volumes to go offline, specify the -force option.

Attention: Do not use the -type fpga option, which upgrades Field Programmable Gate Array (FPGA) firmware, unless directed to do so by an IBM service representative.

13.8 Troubleshooting

This section describes IBM Flex System V7000 error events reporting/notification and understanding the error codes

13.8.1 Using the CMM for troubleshooting tasks

The CMM manages individual nodes within the Flex chassis. For IBM Flex System V7000 Storage Node the CMM manages each canister within an enclosure individually. If you use the CMM to manage the IBM Flex System V7000 you should be aware of the inter-dependencies between all canisters in all the enclosures in the system, and not perform operations that might cause the system as a whole to stop providing access to data. It is always preferable to use the IBM Flex System V7000 management GUI to manage the system that is available. You should only use the CMM when directed to do so by service actions. To view Flex System V7000 Storage Node information using the CMM, use the following steps:

1. Open the CMM and navigate to the chassis map. The chassis map provides a front and rear graphical view of the Flex chassis components. The front view shows compute nodes and storage control enclosures, including canisters and drives. The rear view shows power modules, cooling devices, fan-mux, IO module, and Management module.

   The front view of the chassis map also shows the bay numbering scheme. Generally, bay numbers increment across each row (left to right) from the bottom left to the top. A storage control enclosure takes up four bays. Each canister within a storage control enclosure takes up a bay. Canisters have a slot designation that appears when you hover over them.

   For example, starting at the bottom left, drives in the first enclosure take up bays 1 through 4. The left canister is in bay 1. The right canister is in bay 2. Moving up, the next enclosure takes up bays 5-8. The left canister is in bay 5, and the right canister is in bay 6. The bay numbering scheme continues in the same manner as you move up.

2. Look for colors as an indication of a IBM Flex System V7000 canister's status, as shown in Table 13-12 and Figure 13-19 on page 549

Table 13-12 Colors shown in CMM chassis map

Color	Status
Gray	Normal
Amber	Attention
Red	Error

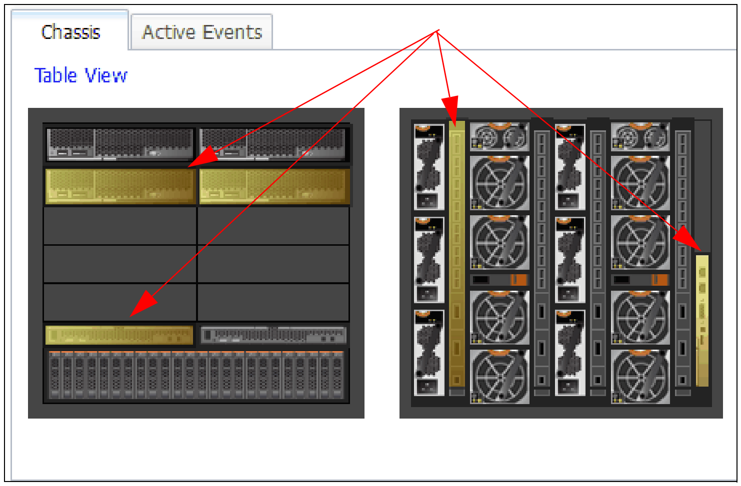


Figure 13-19 CMM chassis map - color status

3. Hover the mouse over a component to view information about it, as shown in Figure 13-20 and Figure 13-21

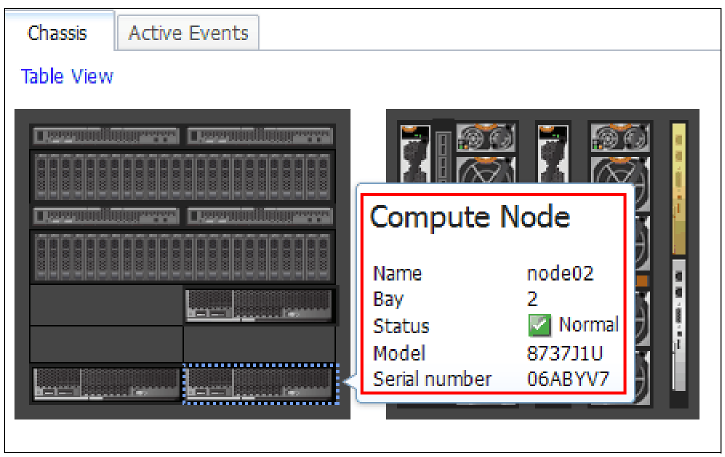


Figure 13-20 CMM chassis map - node canister status information

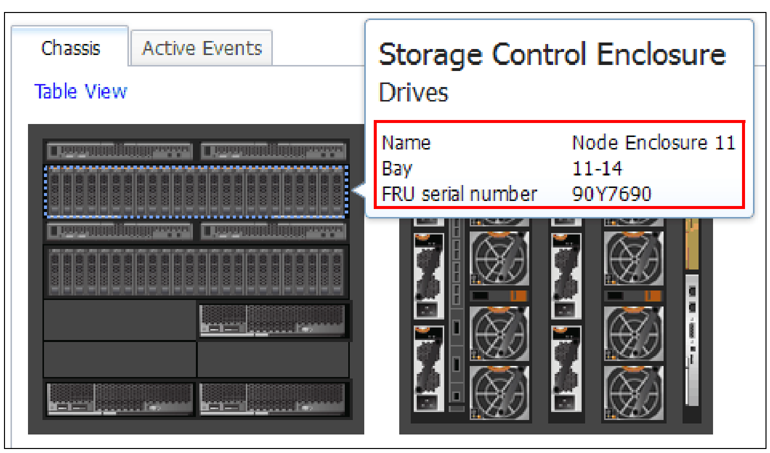


Figure 13-21 CMM chassis map - drive status information

4. Click a component to view detailed information, which appears in tabs in the Details panel, as shown in Figure 13-22 and Figure 13-23.

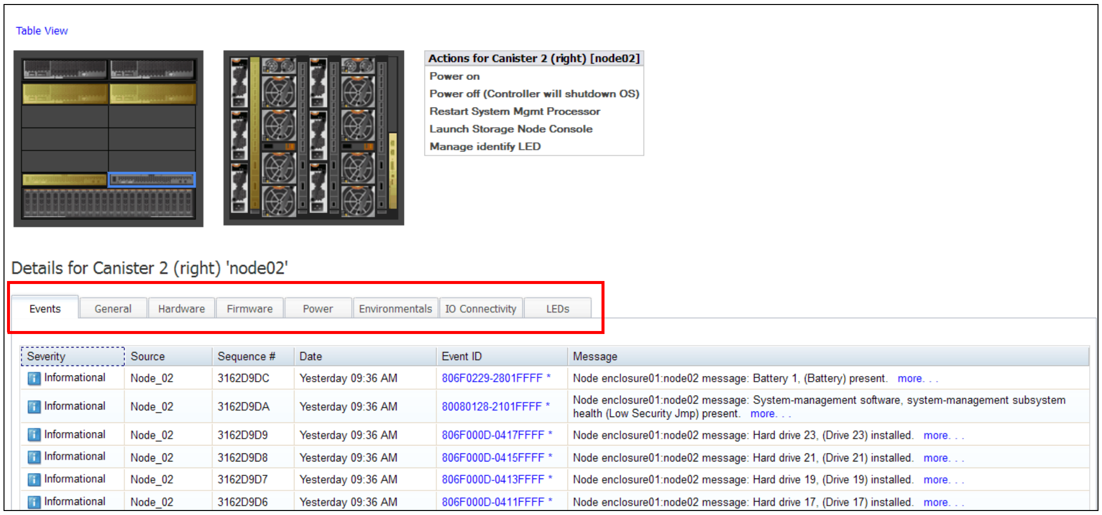


Figure 13-22 CMM chassis map - node canister detail tabs

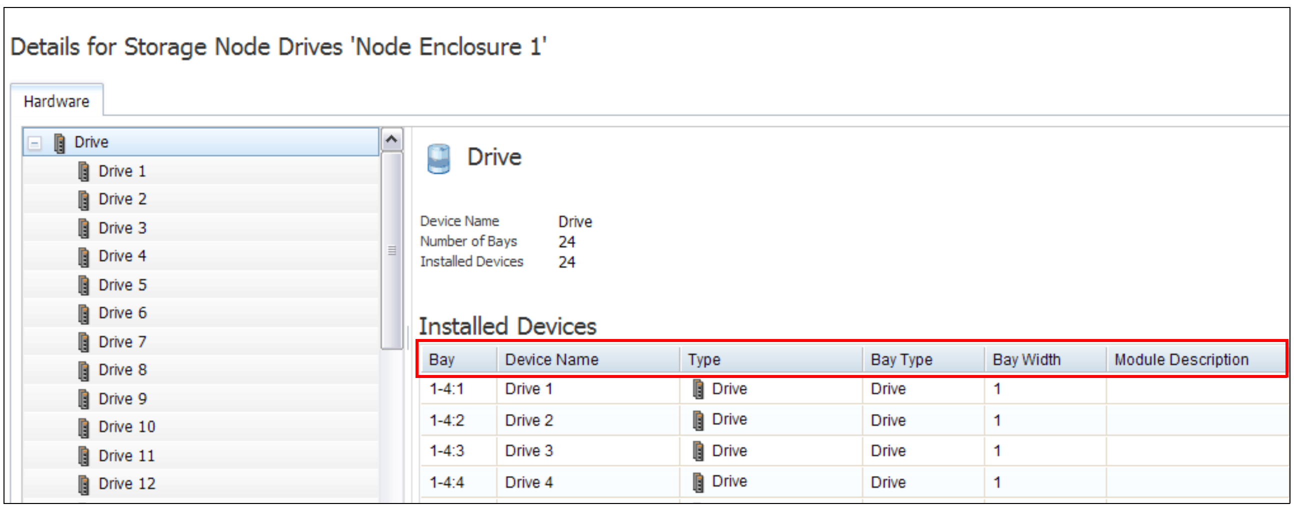


Figure 13-23 CMM chassis map - drive detail tabs

5. Right-click a component to access an action pop-up window, as shown in Figure 13-24 on page 551.

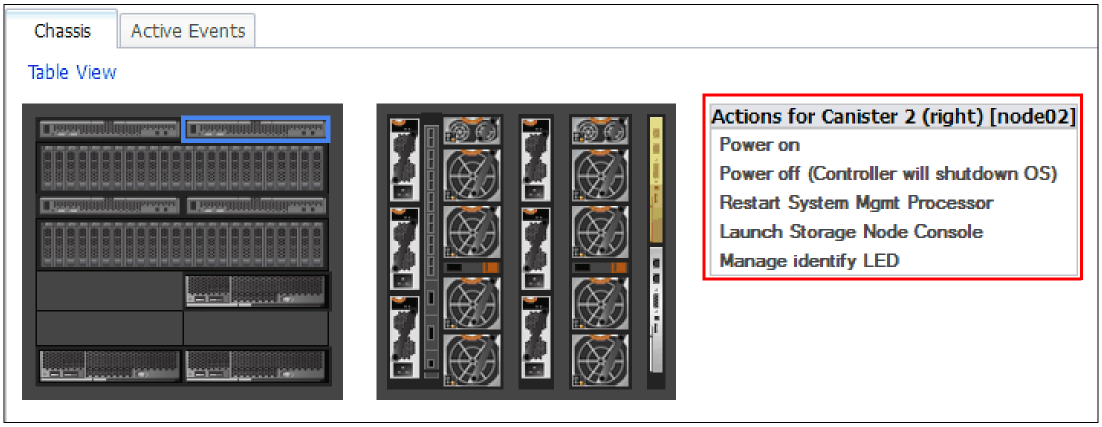


Figure 13-24 CMM chassis map - action pop-up window

> **Attention:** The use of the actions below are described on other pages. Do not use them without referring to those pages.

13.8.2 Using management GUI for troubleshooting tasks

Always use the recommended actions on the Events panel of the management GUI of the Flex System V7000 Storage Node as the starting point to diagnose and resolve a problem. Use the management GUI to run fix procedures topic explains how to run the recommended action for any unfixed alert, and what to do if you cannot run the management GUI, as shown in Figure 13-25.

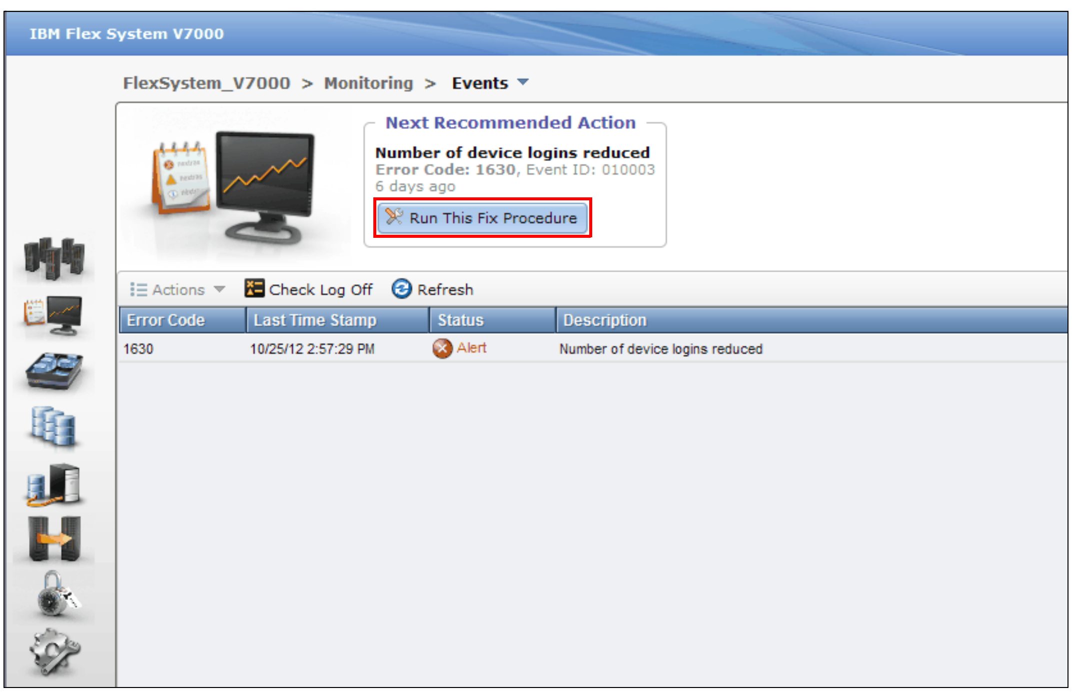


Figure 13-25 Run fix procedures to resolve problems using management GUI

Symptoms of a problem can include hosts being unable to access data, the system sent an e-mail notification, an LED indicates an issue, or the CMM or FSM event logs indicate a

problem. In all these cases the management GUI recommended action is the correct starting point.

For problem that you might experience, that is not resolved by using the management GUI. In those situations, review the symptoms and follow the actions that are provided here. You may be directed to the IBM Flex System V7000 Service Assistant GUI or to Flex System chassis management devices.

It is highly recommended that you set up CLI access to the control enclosure after initial configuration is complete, this allows you to run necessary CLI commands when you cannot start the management GUI or the node canisters in the control enclosure are unable to run the system software.

Attention: After you have created your clustered system, remove hardware components only when directed to do so by the fix procedures. Failure to follow the procedures can result in loss of access to data or loss of data. Follow the fix procedures when servicing a control enclosure.

13.8.3 Removing and replacing parts for troubleshooting and resolving problems

You can remove and replace customer replaceable units (CRUs) from the control enclosure or the expansion enclosure.

Attention: If your system is powered on and performing I/O operations, go to the management GUI and follow the fix procedures. Removing a node without the assistance of the fix procedures can result in loss of data or access to data

Even though many of these procedures are hot-swappable, they are intended to be used only when your system is not up and running and performing I/O operations. If your system is powered on and performing I/O operations, go to the management GUI and follow the fix procedures. Performing the replacement actions without the assistance of the fix procedures can result in loss of data or loss of access to data.

Each replaceable unit has its own removal procedure. Sometimes you can find that a step within a procedure might refer you to a different remove and replace procedure. Complete the new procedure before you continue with the first procedure that you started.

Note: Remove or replace CRU parts only when you are directed to do so. The field replaceable units (FRUs) must be installed only by trained service technicians.

13.8.4 Event reporting

Events that are detected are saved in an event log. As soon as an entry is made in this event log, the condition is analyzed. If any service activity is required, a notification is sent.

The following methods are used to notify you and the IBM Support Center of a new event:

1. If you enabled Simple Network Management Protocol (SNMP), an SNMP trap is sent to an SNMP manager that is configured by the customer.
2. If enabled, log messages can be forwarded from a sender to a receiver on an IP network by using the Syslog protocol.

3. If enabled, event notifications can be forwarded from a sender to a receiver through Call Home email.
4. If Call Home is enabled, critical faults generate a problem management record (PMR) that is sent directly to the appropriate IBM Support Center.

In general, events are classified as either alerts or messages:

- An alert is logged when the event requires some action. Some alerts have an associated error code that defines the service action that is required. The service actions are automated through the fix procedures. If the alert does not have an error code, the alert represents an unexpected change in state. This situation must be investigated to see if it is expected or represents a failure. Investigate an alert and resolve it as soon as it is reported.
- A message is logged when a change that is expected is reported, for instance, an IBM FlashCopy operation completes.

13.8.5 Viewing the event log

Follow the below procedure to view the event log.

On IBM Flex System V7000 Storage node

You can view the event log by using the IBM Flex System V7000 management GUI or the command-line interface (CLI).

You can view the event log by using the **Monitoring** → **Events** options in the management GUI. The event log contains many entries. You can, however, select only the type of information that you need.

You can also view the event log by using the command-line interface (lseventlog). See the “Command-line interface” topic for the command details at the below link

http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/svc_clicommandscontainer_229g0r.html

The event log has a limited size. After it is full, newer entries replace entries that are no longer required. To avoid having a repeated event that fills the event log, some records in the event log refer to multiple occurrences of the same event. When event log entries are coalesced in this way, the time stamp of the first occurrence and the last occurrence of the problem is saved in the log entry. A count of the number of times that the error condition has occurred is also saved in the log entry. Other data refers to the last occurrence of the event.

On IBM FSM

You can select an event filter to display a specific set of events for a selected resource, as shown in Figure 13-26 on page 554 and Figure 13-27 on page 554. You can also select Event Log Preferences to customize how many events to display.

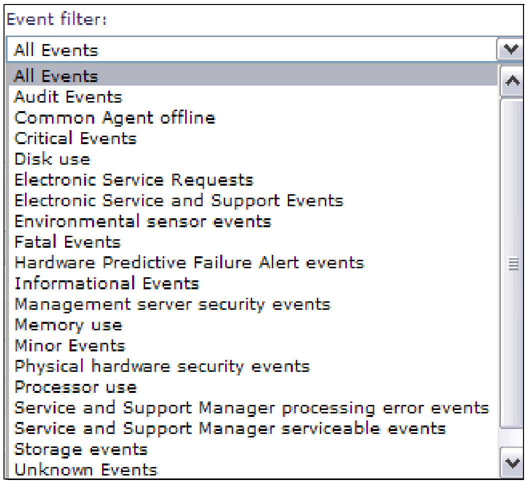


Figure 13-26 Event filters

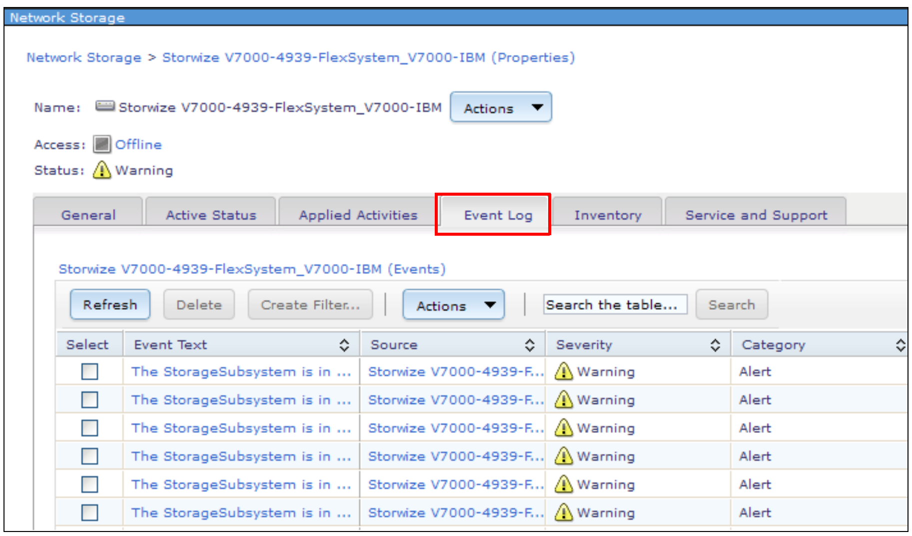


Figure 13-27 Event log collected and displayed for network storage resource

13.8.6 Error event IDs and error codes

Each event ID that requires service has an associated error code. See the IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide PDF on the CD or IBM Flex System Information Center website.for lists of the error event IDs and corresponding error codes.

Node error codes describe failure that relate to a specific node canister. Because node errors are specific to a node, for example, memory has failed, the errors are only reported on that node. However, some of the conditions that the node detects relate to the shared components of the enclosure. In these cases both node canisters in the enclosure report the error.

There are two types of node errors:

- critical node errors
- noncritical node errors

Figure 13-28 shows the navigation window that is used to start the Monitoring wizard for the Viewing System, Viewing System Details, Viewing Events, and See Real Time Performance Graphs.

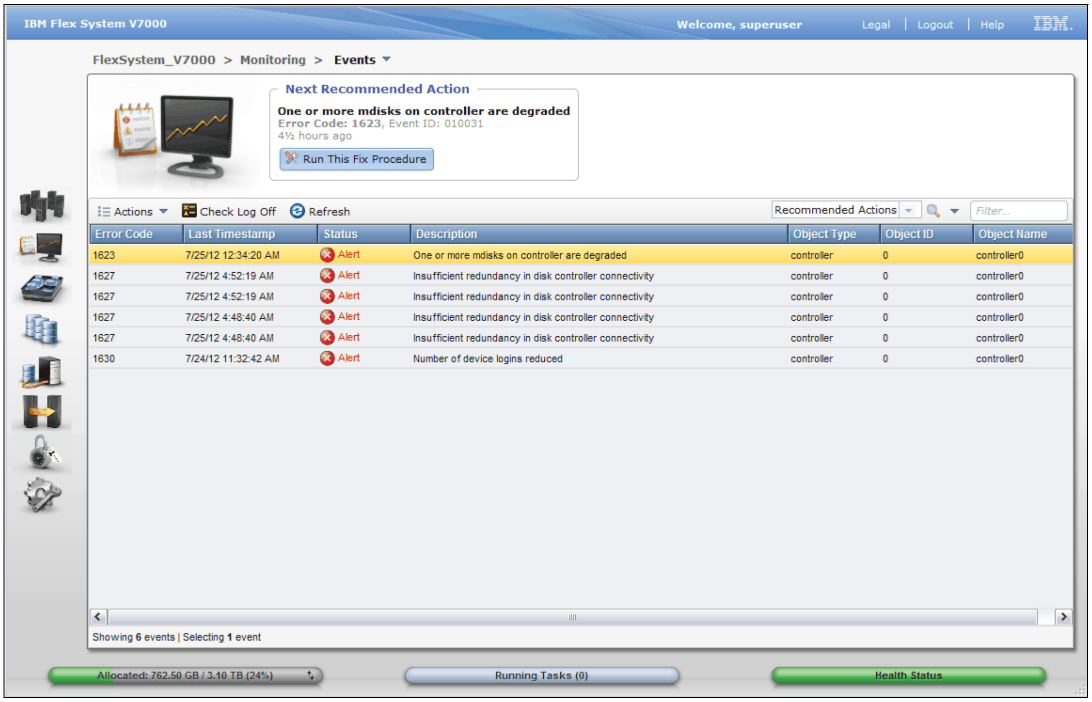


Figure 13-28 Recommended actions

The Recommended Actions tab shows the highest priority maintenance procedure that needs to be run. Use the troubleshooting wizard to allow the IBM Flex System V7000 system to determine the proper order of maintenance procedures. Click **Recommended Actions** → **Run This Fix Procedure** and the highest priority event that needs to be fixed appears (Figure 13-29).

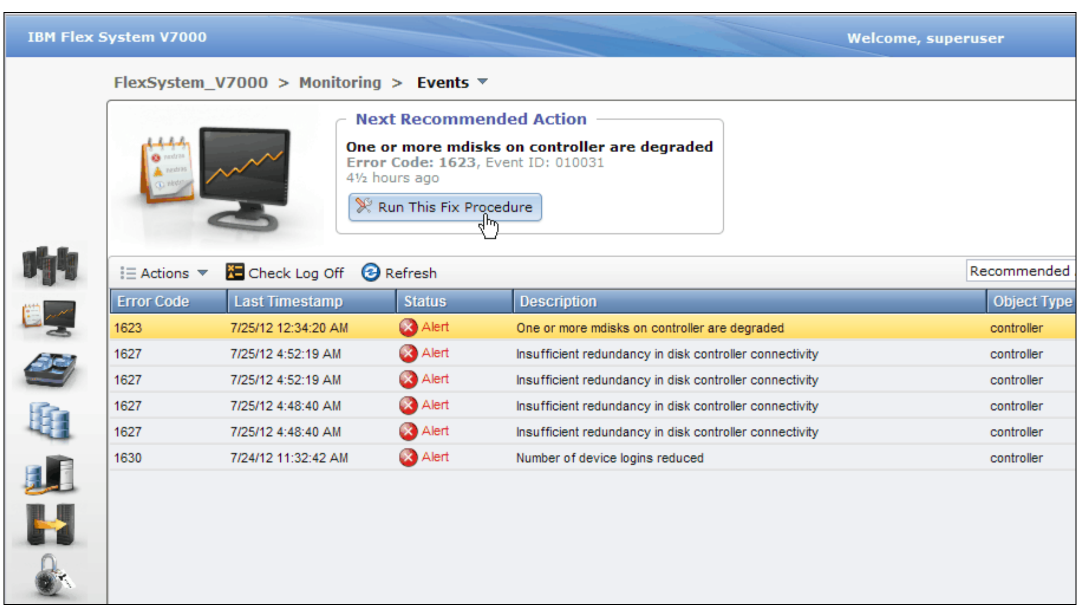


Figure 13-29 Next recommended action

In this example, Number of device logins reduced appears. The next step in this example is to review the physical FC cabling to determine the issue and then click **Run Fix Procedure** to fix the error (Figure 13-30).

A best practice is to review the event logs and recommended actions periodically to ensure that there are no unexpected events, and to configure Call Home so that notification of serious events is done immediately.

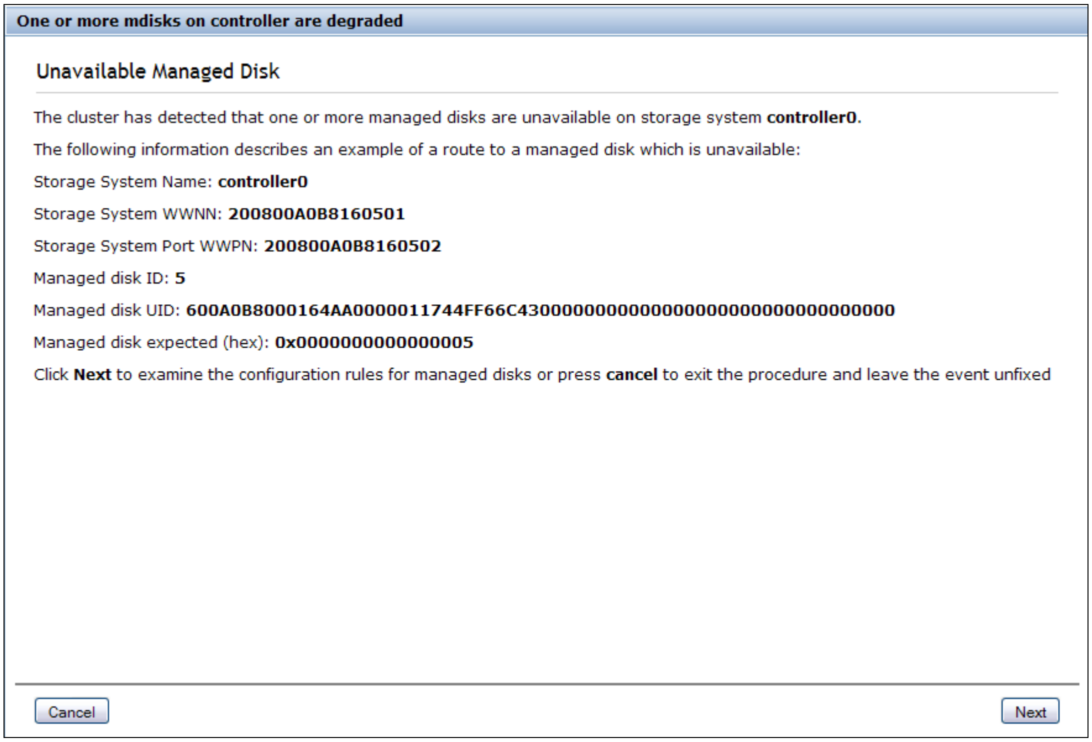


Figure 13-30 Run Fix Procedure

13.9 Audit log navigation

Figure 13-31 shows the Audit Log tab menu item. Click **Audit Log** to view which configuration CLI commands have been run so far on the IBM Flex System V7000 system (information and service commands are not included).

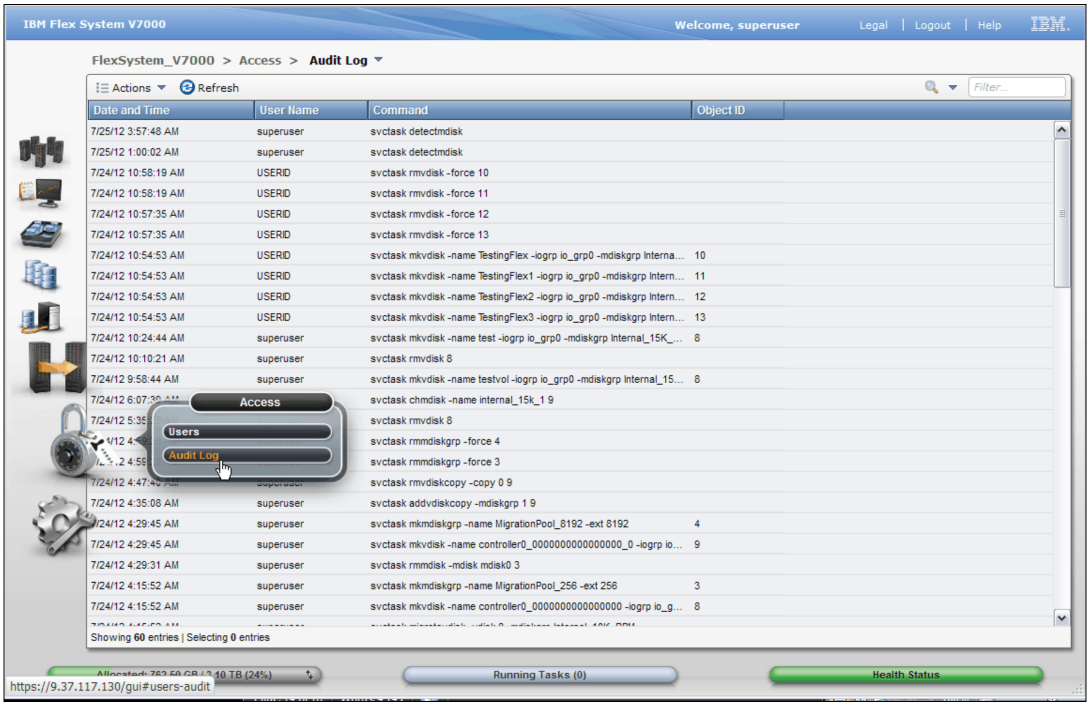


Figure 13-31 Audit Log from Access window

13.10 Support data collection

The following data collection steps may be required when working with IBM Support, as shown in Figure 13-32 on page 558. From this window, click Show full log listing to show all log files currently on IBM Flex System V7000 system.

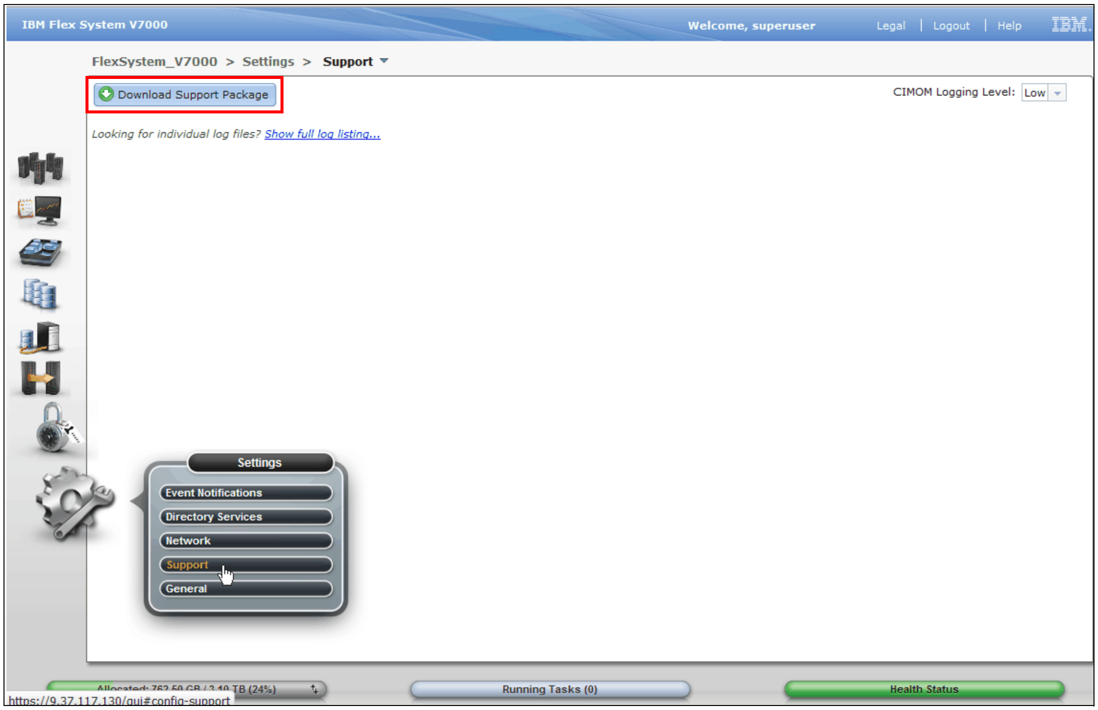


Figure 13-32 Collect support package

Figure 13-33 shows all log files**.** Search for the file you want, Select it, right-click it, and select **Download** to download it to your workstation.

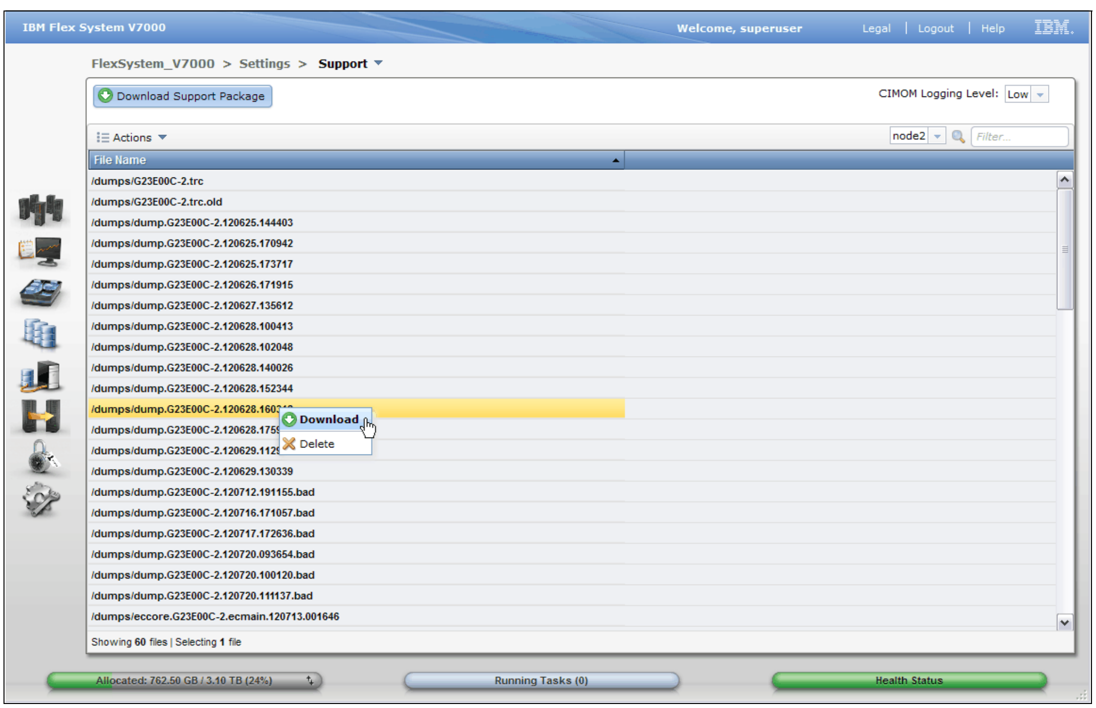


Figure 13-33 Show full log files window

To collect current support package, click download support package and select one of the 4 download options for troubleshooting (Figure 13-34 on page 559) and Save the resulting snap file in a directory you wish.

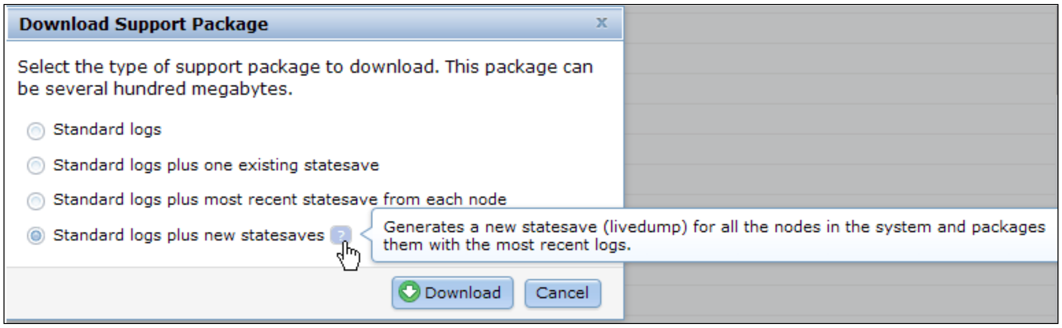


Figure 13-34 Download support package

To upload the resulting snap file to the IBM Support portal after you open a call with IBM support, go to the following address:

`http://www.ecurep.ibm.com/app/upload`

You are ready to call the IBM Support Line or use the IBM Support Portal to open a call. If you use the latter option, go to the following address:

`http://www.ibm.com/support/entry/portal/Open_service_request?brandind=Hardware`

13.10.1 Collecting System Management Server service data using the CMM

Use this procedure when it is necessary to collect System Management Server service data using the CMM.

To collect System Management Server service data using the CMM, use the following steps:

1. Open the CMM. In the menu bar at the top, click Service and Support.
2. In the Service and Support menu, click Download Service Data, and then click Storage Node Service Data.
3. Select the storage node from the Storage Node to download data from: pull-down menu.
4. Select the dump file type from the Dump type: pull-down menu. IBM service requires a service processor dump.
5. Perform the desired dump file action. The Storage Node creates a new dump file on the Storage Node, which overwrites any previous dump files. The CMM transfers the dump file from the storage node to the CMM file system. After the file is transferred, view the /tftproot/service directory using the File Management page. Then, obtain the file using FTP.

   Note: If the file transfer fails, you can view the error details in Not_available.txt, located in /tftproot/service directory.

6. Click OK

13.10.2 Collecting Support files using FSM

Service and Support Manager collects all support files from on this system that have been collected by Service and Support Manager. Support files can contain detailed system information used to help diagnose a serviceable hardware problem, dump files collected from an endpoint system, event logs, and more. Use this view to see more details about collected

support files, to manually collect additional support files, and to submit collected files to the IBM service provider for analysis, as show in Figure 13-35.

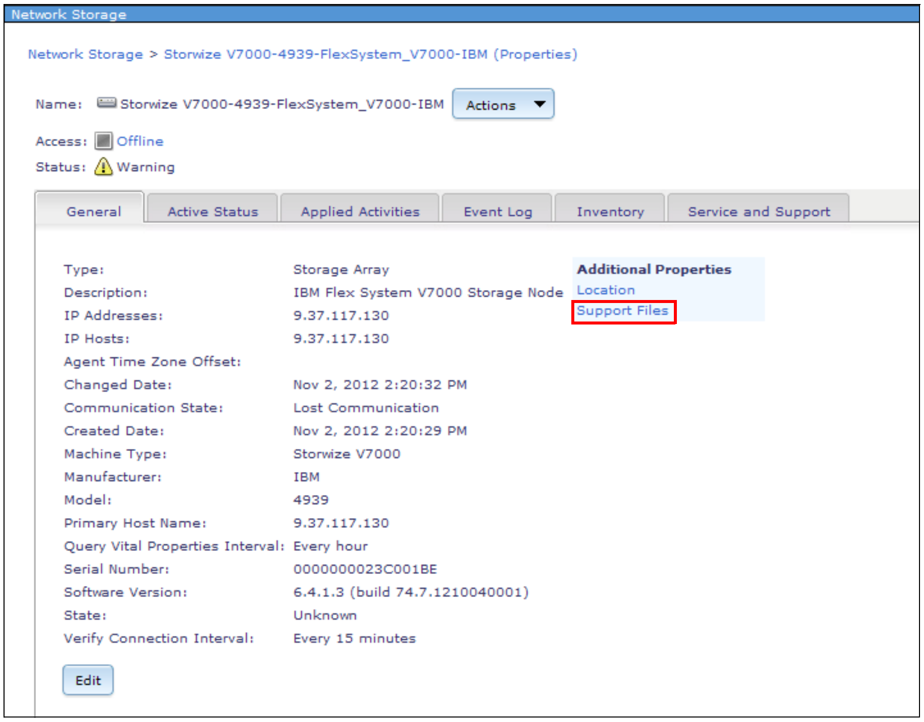


Figure 13-35 FSM collects support files

The following view displays all support files on this system that have been collected by Service and Support Manager. Support files can contain detailed system information used to help diagnose a serviceable hardware problem, dump files collected from an endpoint system, event logs, and more. Use this view to see more details about collected support files, to manually collect additional support files, and to submit collected files to the IBM service provider for analysis, as shown in Figure 13-36 on page 561.

Support files associated with a problem cannot be submitted unless the problem itself has been submitted to IBM support and is currently in a submitted state. Click on Problems to view the current status of the problem and ensure that the status is submitted before attempting to submit any associated support files,

Note: Support files can only be collected on systems in the monitored group.

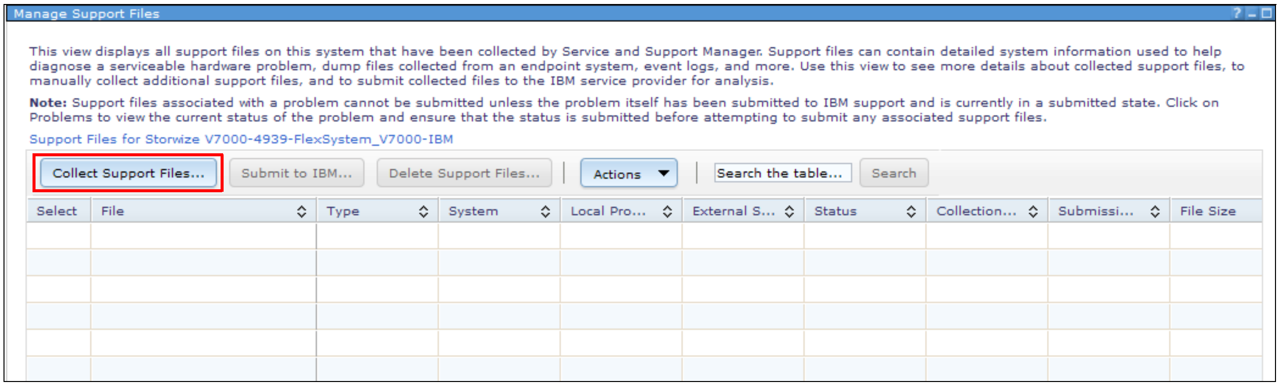


Figure 13-36 FSM Support files collection

Support files associated with a problem cannot be submitted unless the problem itself has been submitted to IBM support and is currently in a submitted state. These Support files can only be collected on systems in the monitored group. See the IBM Flex System Manager Node Infocenter (link below) and *Implementing Systems Management of IBM PureFlex System*, SG24-8060 for more information.

`http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_product_page.html`

13.11 Using event notifications

IBM Flex System V7000 Storage Node can use Simple Network Management Protocol (SNMP) traps, syslog messages, and a Call Home email to notify you and the IBM Support Center when significant events are detected. Any combination of these notification methods can be used simultaneously.

Each event that IBM Flex System V7000 detects is assigned a notification type of Error, Warning, or Information. You can configure IBM Flex System V7000 to send each type of notification to specific recipients.

Correct any issues reported by your system as soon as possible. To avoid monitoring for new events by constantly monitoring the management GUI, configure your system to send notifications when a new event is reported. Select the type of event that you want to be notified about. For example, restrict notifications to just events that require immediate action. Several event notification mechanisms exist:

- SNMP traps

  SNMP is a standard protocol for managing networks and exchanging messages. IBM Flex System V7000 can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that IBM Flex System V7000 sends. You can use the management GUI or the IBM Flex System V7000 command-line interface to configure and modify your SNMP settings. You can use the Management Information Base (MIB) file for SNMP to configure a network management program to receive SNMP messages that are sent by the IBM Flex System V7000. This file can be used with SNMP messages from all versions of IBM Flex System V7000 Software.,

- Syslog messages

  The syslog protocol is a standard protocol for forwarding log messages from a sender to a receiver on an IP network. The IP network can be either IPv4 or IPv6. IBM Flex System

V7000 can send syslog messages that notify personnel about an event. IBM Flex System V7000 can transmit syslog messages in either expanded or concise format. You can use a syslog manager to view the syslog messages that IBM Flex System V7000 sends. IBM Flex System V7000 uses the User Datagram Protocol (UDP) to transmit the syslog message. You can use the management GUI or the IBM Flex System V7000 command-line interface to configure and modify your syslog settings

- Email

  An event notification can be sent to one or more email addresses. This mechanism notifies individuals of problems. Individuals can receive notifications wherever they have email access which includes mobile devices.

13.12 Configuring Call Home

The Call Home feature transmits operational and error-related data to you and IBM through a Simple Mail Transfer Protocol (SMTP) server connection in the form of an event notification email. When configured, this function alerts IBM service personnel about hardware failures and potentially serious configuration or environmental issues. You can use the call home function if you have a maintenance contract with IBM or if the IBM Flex System V7000 is within the warranty period.

To send email, you must configure at least one SMTP server. You can specify as many as five additional SMTP servers for backup purposes. The SMTP server must accept the relaying of email from the IBM Flex System V7000 clustered system IP address. You can then use the management GUI or the IBM Flex System V7000 command-line interface to configure the email settings, including contact information and email recipients. Set the reply address to a valid email address. Send a test email to check that all connections and infrastructure areset up correctly.

> **Note:** Enable the Call Home event notification mechanism is highly recommended. However, you can disable the Call Home function at any time using the management GUI or the IBM Flex System V7000 command-line interface.
>
> If you set up Call Home to IBM, ensure that the contact details that you configure are correct and kept up to date as personnel change.

You can configure Call Home differently, depending on whether IBM Flex System V7000 includes FSM or not.

13.12.1 Configuring Call Home if FSM is not included

You configure Call Home on the IBM Flex System V7000 by enabling Email Event Notification via the management GUI.

You also configure Call Home on CMM to cover chassis related problems, which are relevant to IBM Flex System V7000 operation. CMM performs Call Home for problems with IBM Flex System V7000 canisters, including if they are dead and cannot boot, or if they have a bad DIMM.

Follow this procedure to configure the call home.

1. If the step to configure Support Notifications (Step 5 of 7) was complete already during initial setup, as shown in Figure 13-37 on page 563 and Figure 13-38 on page 563,

besides sending a test email to check that all connections and infrastructure are set up correctly, no further Call Home configuration is required.

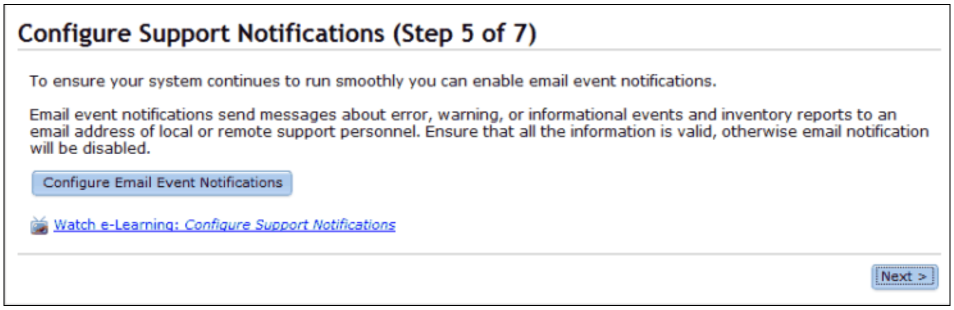


Figure 13-37 Configure email event notification during initial setup

2. To enable the Call Home event notification function, go to IBM Flex system V7000 management **GUI** → **Settings** → **Event Notifications** → **Email** → **Enable Email Event Notification**, as shown in Figure 13-38.

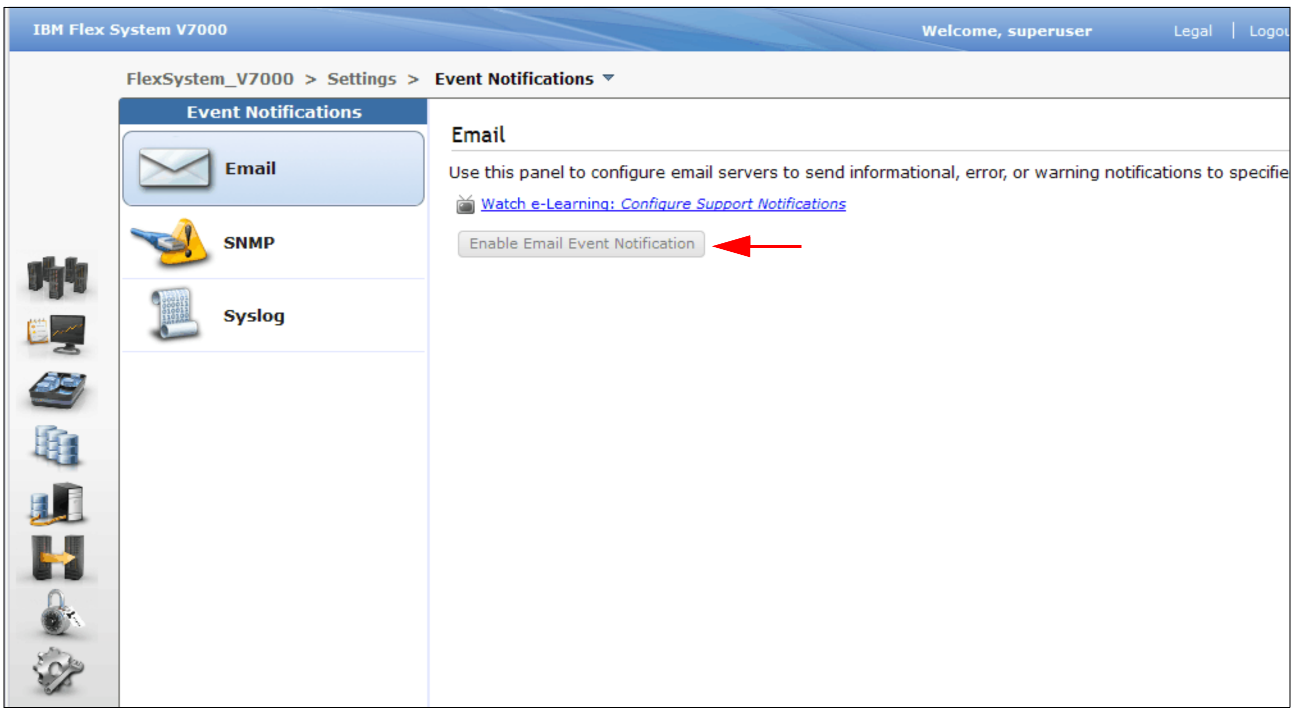


Figure 13-38 Enable email event notification after the initial setup is complete

3. Follow the instruction to complete the configuration, as shown in Figure 13-39.

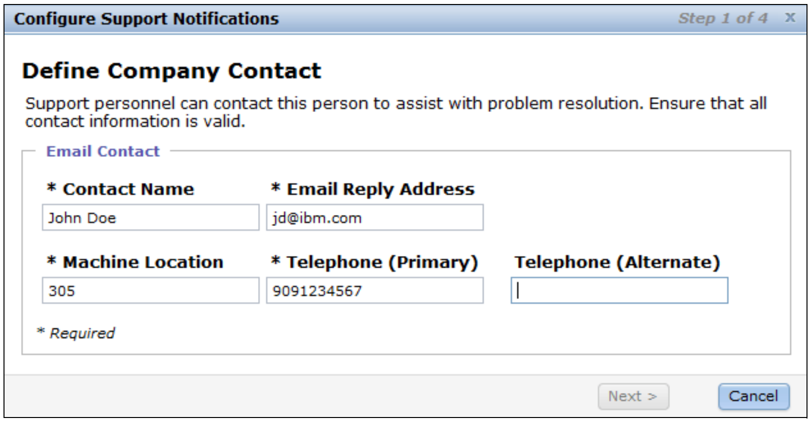


Figure 13-39 Define company contact information

13.12.2 Configuring Call Home if FSM is included

Note: All the topics mentioned in the procedure below are available for reference on the IBM Flex System Information Center at the below link:

`http://publib.boulder.ibm.com/infocenter/flexsys/information/index.jsp`

You configure Call Home on FSM. FSM automatically takes control of Call Home. See ***Enabling a system for monitoring by the Service and Support Manager*** for information on enabling system monitoring (including performing a system discovery, requesting access to a secured system, and collecting inventory). See ***Event actions*** for general information on event actions, and ***Configure e-mail notification for Service and Support Manager using automation plans*** for instructions on configuring e-mail notification for Service and Support Manager using automation plans.

If you launch the IBM Flex System V7000 installation sequence from FSM, you will not see Call Home configuration prompts in the Setup wizard. You do not perform Call Home configuration tasks on the IBM Flex System V7000.

FSM performs Call Home for problems that IBM Flex system V7000 and CMM identify. They appear in Service and Support Manager. See ***Setting up Electronic Service Agent™ (ESA)*** for additional information.

IBM Service and Support Manager monitors, tracks, and captures system hardware errors and service information and reports serviceable problems directly to IBM Support using the IBM Electronic Service Agent, which is integrated with the IBM Flex System Manager to securely transmit serviceable hardware problems and associated support files to IBM Support.

When Electronic Service Agent is enabled and the IBM Flex System Manager receives a serviceable hardware problem from a managed device, the IBM Flex System Manager collects hardware problem data from the managed endpoint and stores this data in the form of a support file. The actual contents of the support file is predetermined based on the type of managed endpoint for which the data is collected.

After the support file has been collected, the IBM Flex System Manager securely transmits any serviceable hardware problems and associated support files to IBM Support. IBM Support analyzes this data and contacts the customer for problem resolution.

Follow the below procedure to configure call home if the FSM is included.

1. You can set up and enable Electronic Service Agent from the Additional Setup tab of the IBM Flex System Manager home page Getting started with Electronic Service Agent welcome screen, as shown in Figure 13-40 on page 565.

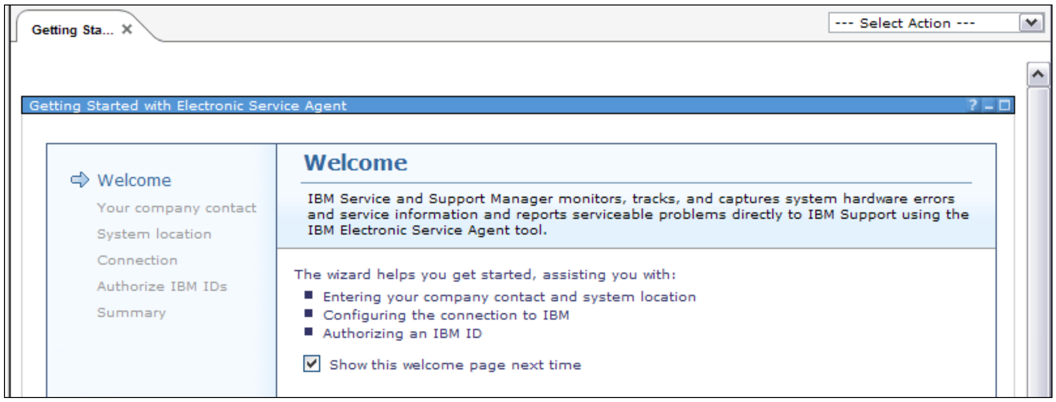


Figure 13-40 Getting Started with Electronic Service Agent welcome screen

2. Click Next, Provide information about the person that IBM Support may contact about a problem reported by Electronic Service Agent. Figure 13-41.

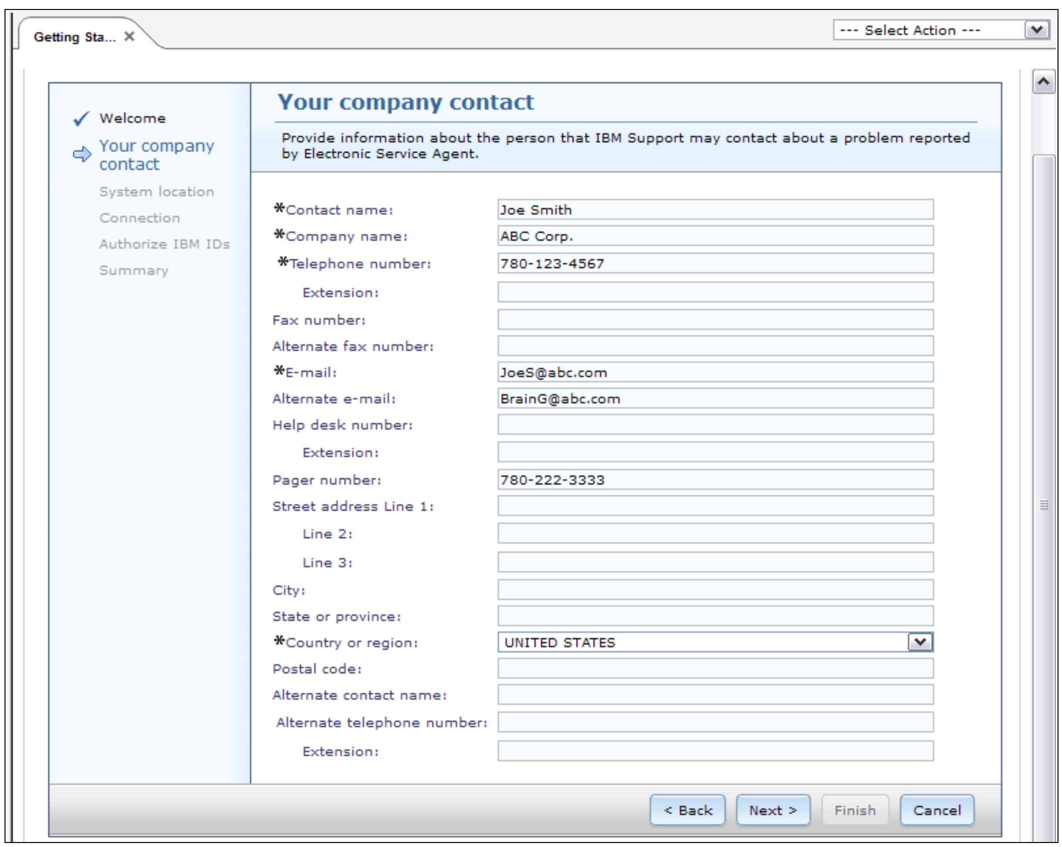


Figure 13-41 Company contact information

3. Click Next. Provide default information about the physical locations of your systems. Information can be overridden for specific systems by clicking Resource Explorer,

selecting a system, and clicking Location under the Additional Properties heading, as shown in Figure 13-42 on page 566.

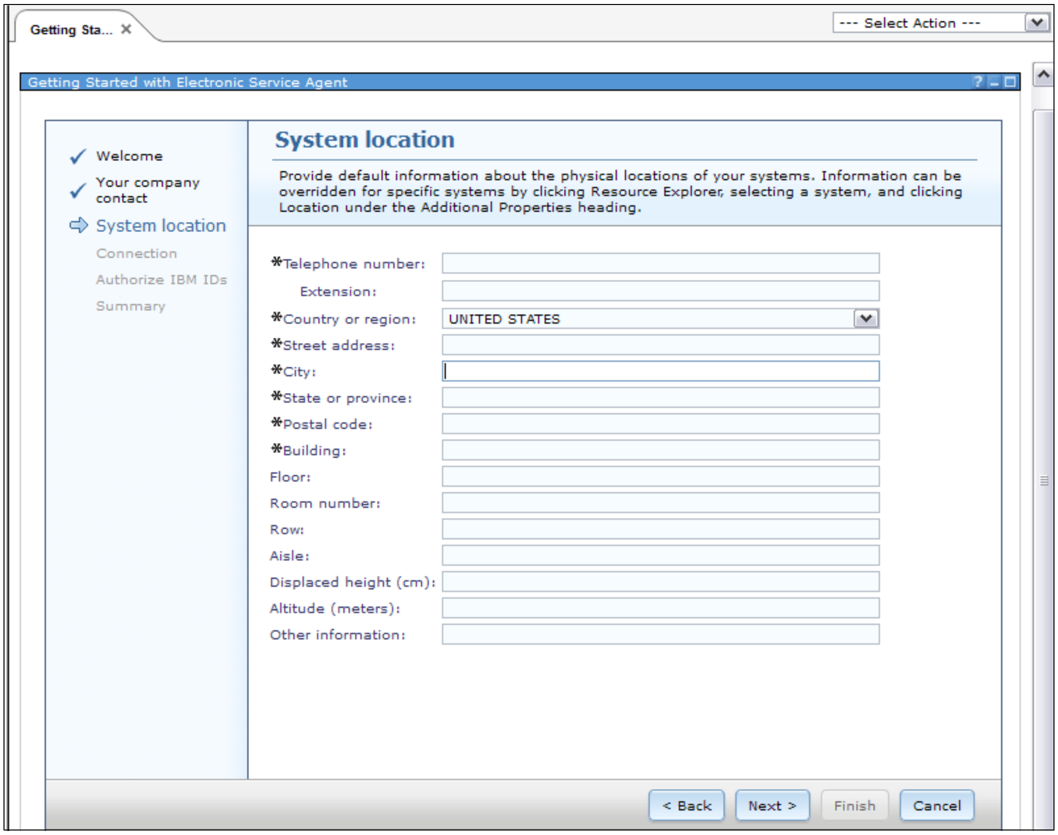


Figure 13-42 System location

4. Click OK, An Internet connection is required to use this function. Specify how the Internet should be accessed. Click Test Internet Connection to verify your internet access is working properly, as shown in Figure 6 on page 568.

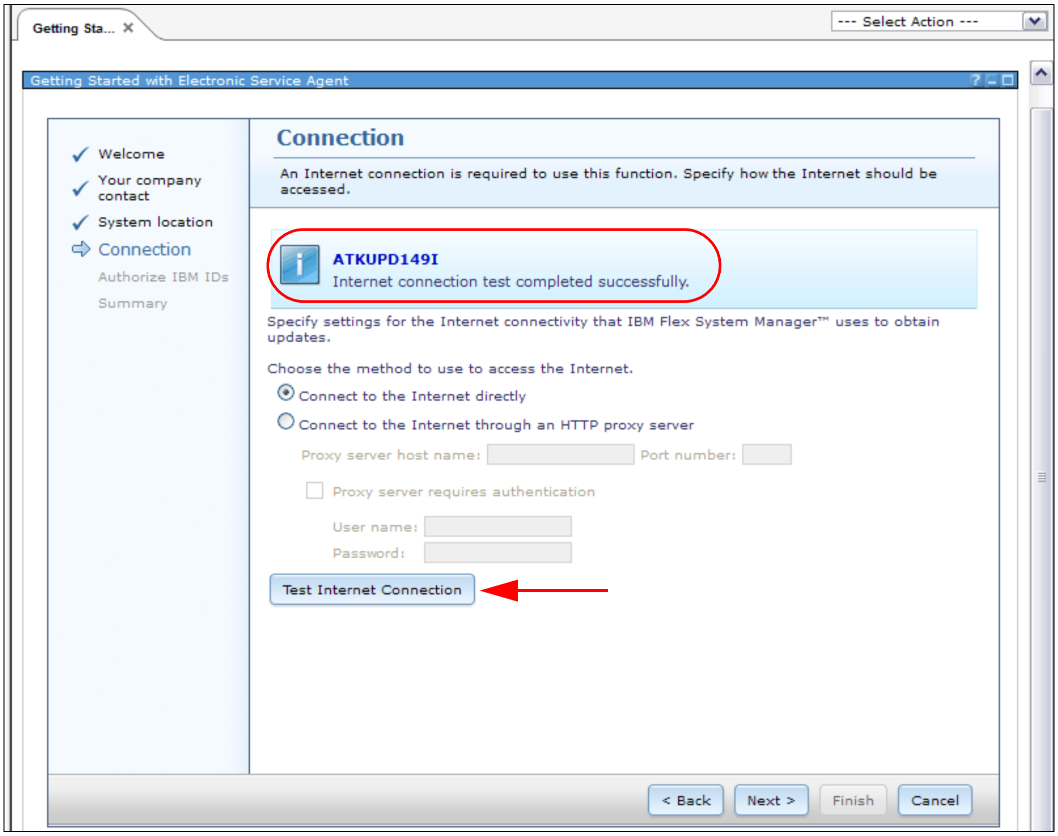


Figure 13-43 Select internet access

5. Click Next. Provide an IBM ID to be associated with information sent by Electronic Service Agent, as shown in Figure 6 on page 568.

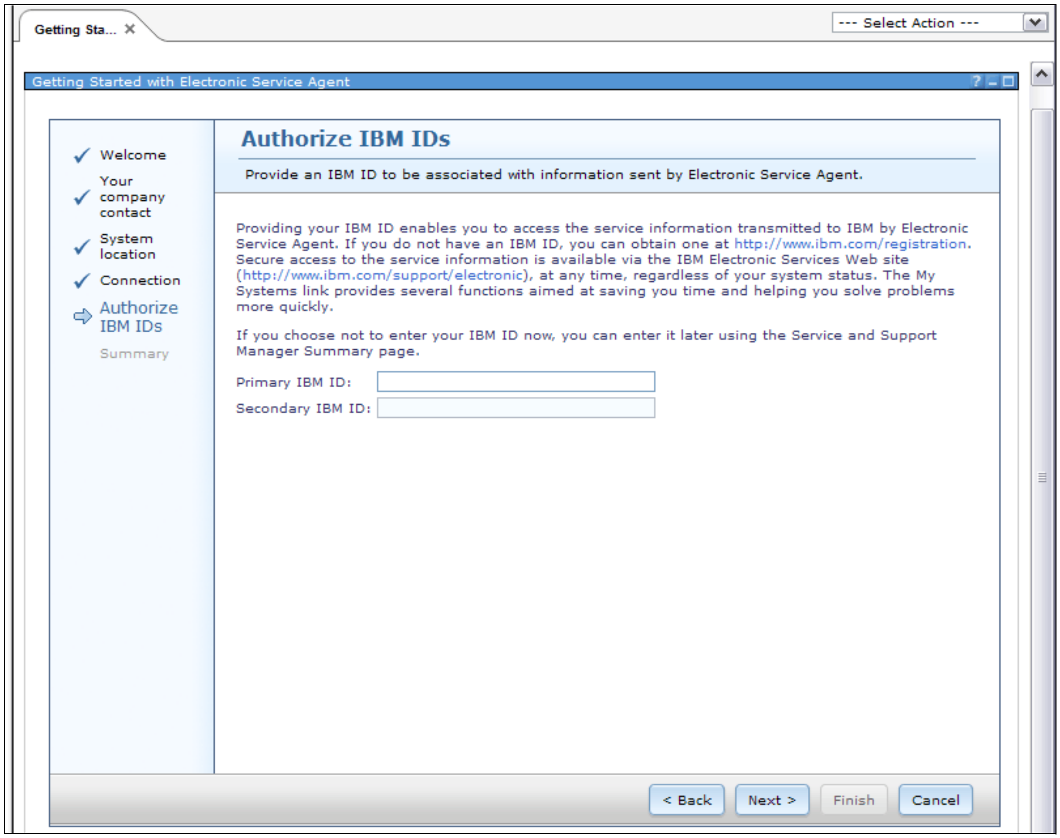


Figure 13-44 Provide IBM ID access

6. Click Next, Review the Summary page and click Finish if all information displayed is correct, as shown in Figure 13-45 on page 569.

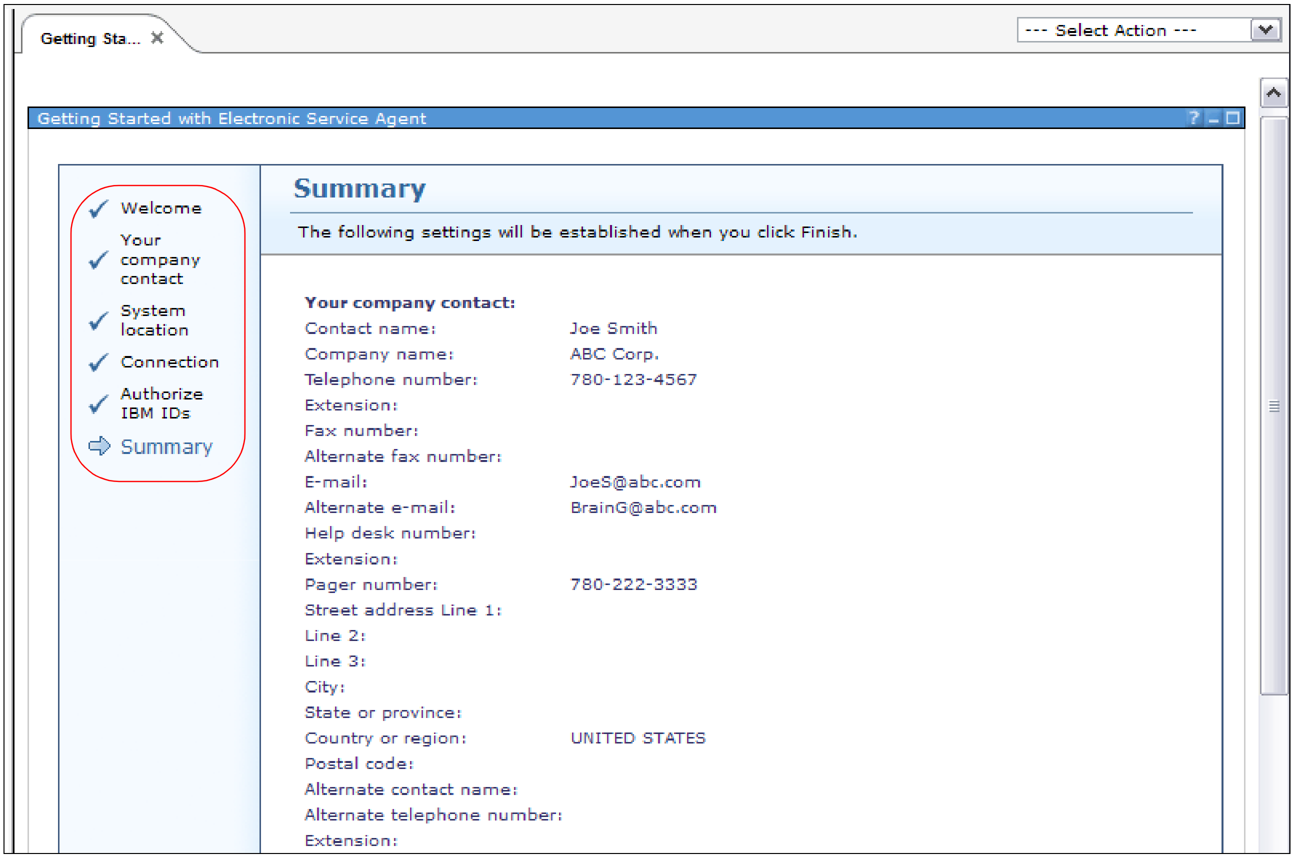


Figure 13-45 Summary screen for review

13.13 IBM Flex System V7000 Storage Node Power On and Off

You may shut down only one node canister using CMM or the entire cluster using IBM Flex System V7000 management GUI. When shutting down only one node canister, all the activities remain active on the partner node; however, when you shut down the entire cluster, your access to volumes and data will be lost until you manually restart the system.

Follow these general procedures to power on or to power off the system. The procedures must be followed in the order given. See << Author comment - add a reference to InfoCenter when available>>for more details.

13.13.1 Powering on the system

Follow the below procedure to power on the system.

1. If the chassis is powered-off, physically go to the chassis and press the power switch. Normally, this automatically causes power on to the Flex System V7000 Storage Node.
2. If the Flex System V7000 Storage Node does not automatically power on, use the CMM or FSM GUI / CLI to request power-on for the Flex System V7000 Storage Node.
3. Power on expansion enclosures before control enclosures

13.13.2 Powering off the system

Follow the below procedure to power off the system.

1. Stop any hosts using the Flex System V7000 Storage Node.
2. Use the Flex System V7000 Storage Node management GUI to shut down the system, as shown in Figure 13-46.

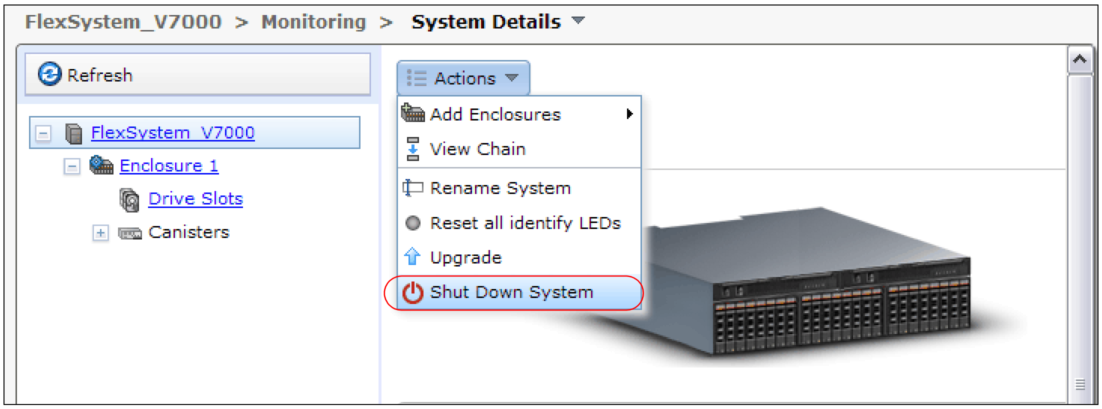


Figure 13-46 Shutting down the system from IBM Flex System V7000 GUI

Attention: Ensure that you have stopped all FlashCopy mappings, Remote Copy relationships, data migration operations, and forced deletions before continuing. Click **Yes** to begin the shutdown process, as shown in Figure 13-47.

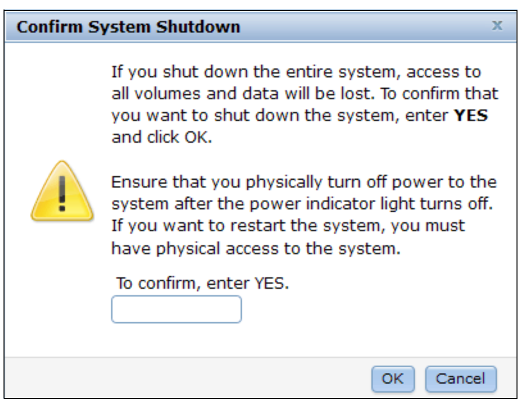


Figure 13-47 Confirmation to start the shutdown process

13.13.3 Shutting down using IBM Flex System V7000 command-line interface

You can also shut down an IBM Flex System V7000 Storage Node by using the CLI and the PuTTY utility. Run the ***stopcluster*** command to shut down the system:

13.13.4 Powering off a node using the CMM

Use this procedure when it is necessary to power off the Flex System V7000 Storage Node using the CMM. Make sure you select the correct node canister before proceeding. Power off

a Flex System V7000 Storage Node canister using the Flex System V7000 Storage Node service GUI.

1. Open the CMM and navigate to the chassis map. The chassis map displayed by the CMM provides a graphical view of the IBM Flex System Enterprise Chassis components. Using the mouse, locate the node that you want to power off, as shown in Figure 13-48.

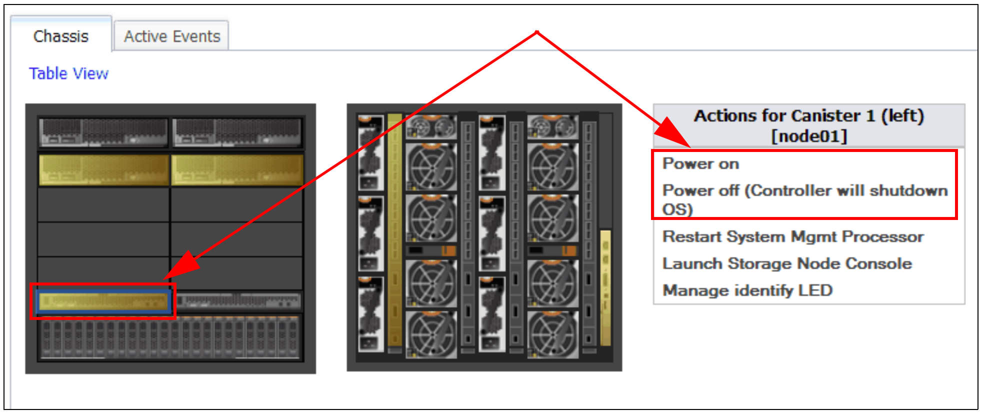


Figure 13-48 MM chassis map - powering on and off individual node canister

> **Attention:** This operation can result in permanent loss of data on storage volumes or loss of access to the volumes, Use the storage management user interface provided by the storage node to power off the storage system or individual storage controllers. Only continue if the service procedures have directed the use of this CMM power control operation.

2. Click the node to display the Actions menu.
3. In the Actions menu click Power Off. (Controller will shutdown OS).

13.13.5 Powering on a node using the CMM

Use this procedure when it is necessary to power on the IBM Flex System V7000 Storage Node using the CMM, as shown in Figure 13-48.

1. Open the CMM and navigate to the chassis map. The chassis map displayed by the CMM provides a graphical view of the Flex chassis components.
2. Using the mouse, locate the node that you want to power on.
3. Click the node to display the Actions menu.
4. In the Actions menu click Power On.

A

CLI Setup and Configuration

This section describes the setup of the CLI and how to setup the host for an Secure Shell connection. To demonstrate this capability we used PuTTY, a freeware tool that is available at the following url:

Command-line interface

The IBM Flex System V7000 Storage Node also has a powerful CLI, which offers even more functions than those provided by the GUI. This section is not intended to be a detailed guide to the CLI, as that topic is beyond the scope of this book. The basic configuration of the IBM Flex System V7000 Storage Node CLI and some example commands are covered. However, the CLI commands are the same as in the SAN Volume Controller, and in addition, there are more commands that are available to manage internal storage. If a task completes in the GUI, the CLI command is always displayed in the details, as shown throughout this book.

Detailed CLI information for the IBM Flex System V7000 Storage Node is available in the *Command-Line Interface User's Guide* which can be found in the Information Center for the IBM Storwize V7000 product, which can be found at the following address:

http://publib.boulder.ibm.com/infocenter/flexsys/information/index.jsp

Implementing the IBM System Storage SAN Volume Controller V6.3, SG24-7933 also has a great deal of information about using the CLI, and the command examples in that book also apply to the IBM Flex System V7000 Storage Node as well.

Basic setup

In the IBM Flex System V7000 Storage Node GUI, authentication is done by using a user name and a password. The CLI uses a Secure Shell to connect from the host to the IBM Flex System V7000 Storage Node. To access either a private and a public key pair or user name and password is necessary. The following steps are required to enable CLI access with an SSH session which would likely be used with a *PuTTY SSH connection* which is commonly for connections and uses SSH keys:

- A public key and a private key are generated together as a pair.
- A public key is uploaded to the IBM Flex System V7000 Storage Node system through the GUI.
- A client Secure Shell (SSH) tool must be configured to authenticate with the private key.
- A secure connection can be established between the client and the IBM Flex System V7000 Storage Node.

Secure Shell is the communication vehicle between the management workstation and the IBM Flex System V7000 Storage Node. The SSH client provides a secure environment from which to connect to a remote machine. It uses the principles of public and private keys for authentication.

SSH keys are generated by the SSH client software. The SSH keys include a public key, which is uploaded and maintained by the cluster, and a private key, which is kept private on the workstation that is running the SSH client. These keys authorize specific users to access the administration and service functions on the system. Each key pair is associated with a user-defined ID string that can consist of up to 40 characters. Up to 100 keys can be stored on the system. New IDs and keys can be added, and unwanted IDs and keys can be deleted. To use the CLI, an SSH client must be installed on that system, the SSH key pair must be generated on the client system, and the client's SSH public key must be stored on the IBM Flex System V7000 Storage Node.

The SSH client used in this book is PuTTY. There is also a PuTTY key generator that can be used to generate the private and public key pair. The PuTTY client can be downloaded at no cost at the following address:

`http://www.chiark.greenend.org.uk`

Download the following tools:

- PuTTY SSH client: `putty.exe`
- PuTTY key generator: `puttygen.exe`

Generating a public and private key pair

To generate a public and private key pair, complete the following steps:

1. Start the PuTTY key generator to generate the public and private key pair (Figure A-1).

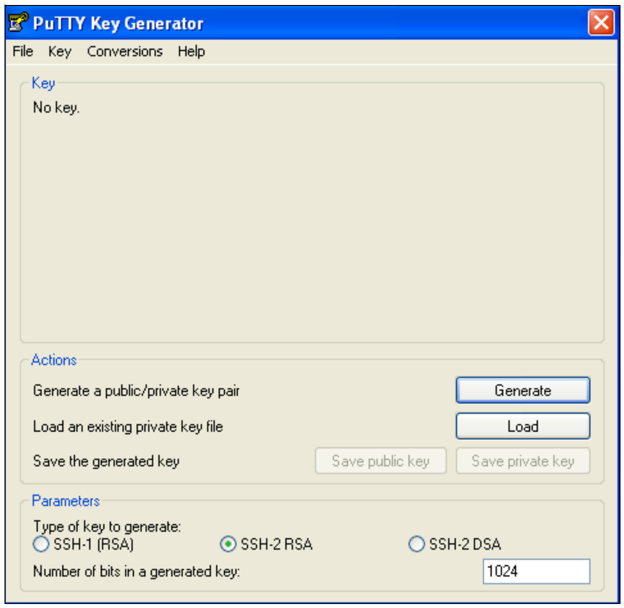


Figure A-1 PuTTY key generator

   Make sure that the following options are selected:

   – SSH2 RSA
   – Number of bits in a generated key: 1024

2. Click **Generate** and move the cursor over the blank area to generate the keys (Figure A-2).

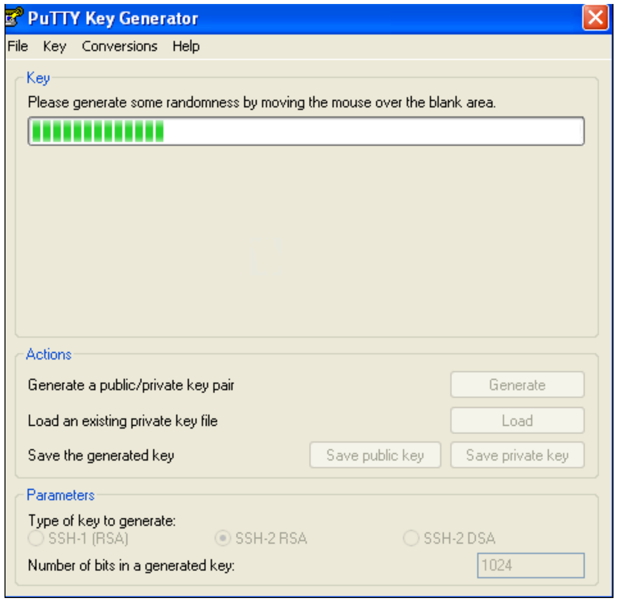


Figure A-2 Generate keys

To generate keys: The blank area indicated by the message is the large blank rectangle on the GUI inside the section of the GUI labeled Key. Continue to move the mouse pointer over the blank area until the progress bar reaches the far right. This action generates random characters to create a unique key pair.

3. After the keys are generated, save them for later use. Click **Save public key** (Figure A-3).

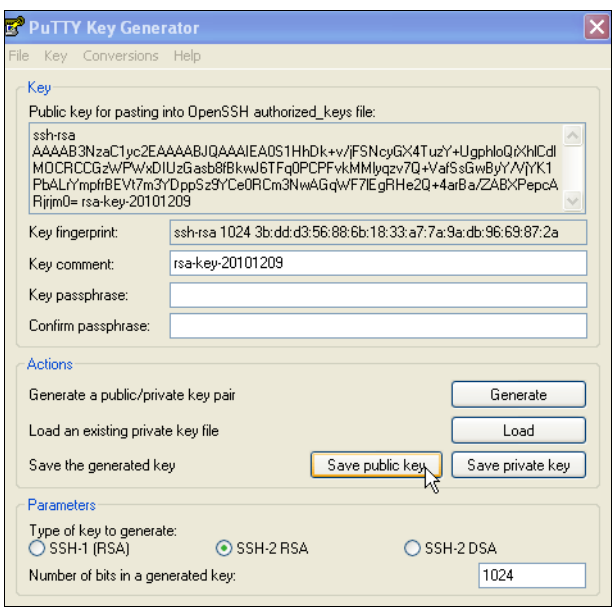


Figure A-3 Save public key

4. You are prompted for a name (for example, pubkey) and a location for the public key (for example, `C:\Support Utils\PuTTY`). Click **Save**.

   Ensure that you record the name and location, because the name and location of this SSH public key must be specified later.

   Public key extension: By default, the PuTTY key generator saves the public key with no extension. Use the string “pub” for naming the public key, for example, “pubkey”, to easily differentiate the SSH public key from the SSH private key.

5. Click **Save private key** (Figure A-4).

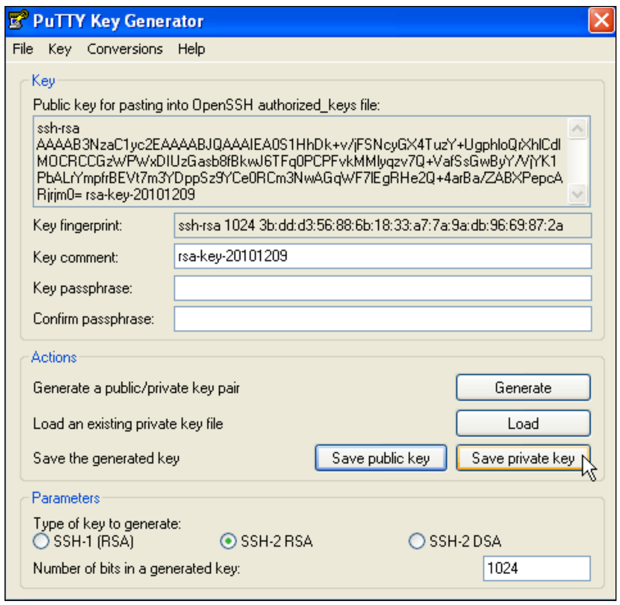


Figure A-4 Save private key

6. You are prompted with a warning message (Figure A-5). Click **Yes** to save the private key without a passphrase.

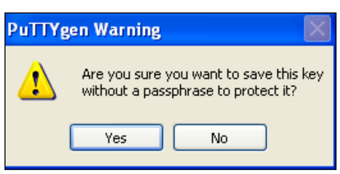


Figure A-5 Confirm the security warning

7. When prompted, enter a name (for example, “icat”), select a secure place as the location, and click **Save**.

   Key generator: The PuTTY key generator saves the private key with the PPK extension.

8. Close the PuTTY key generator.

Uploading SSH public key to the IBM Flex System V7000 Storage Node

After you have created your SSH key pair, you need to upload your SSH public key onto the IBM Flex System V7000 Storage Node. Complete the following steps:

1. Open the user section (Figure 13-49).

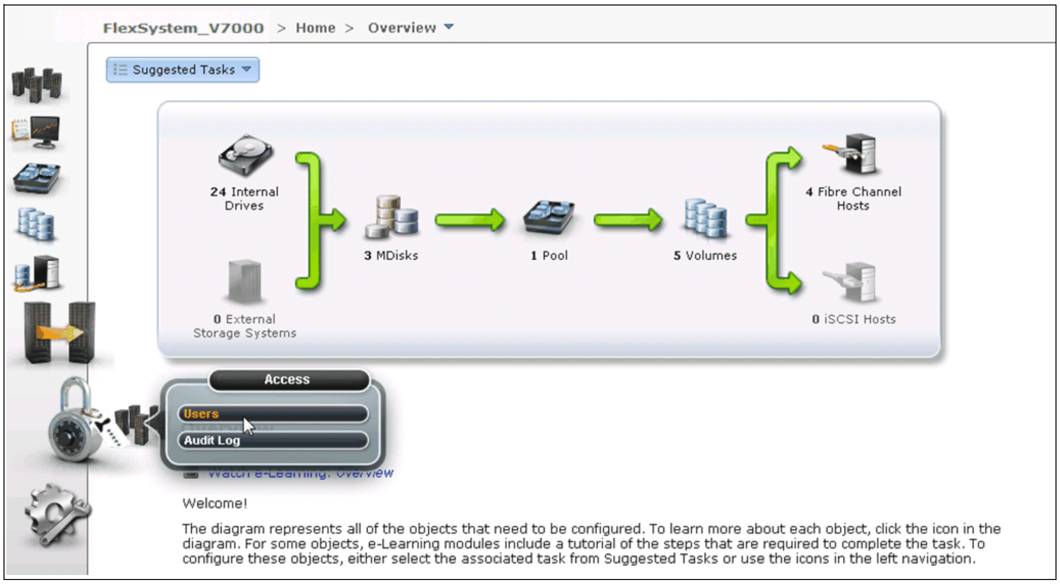


Figure 13-49 Open Users section

2. Right-click the user for which you want to upload the key and click **Properties** (Figure A-6).

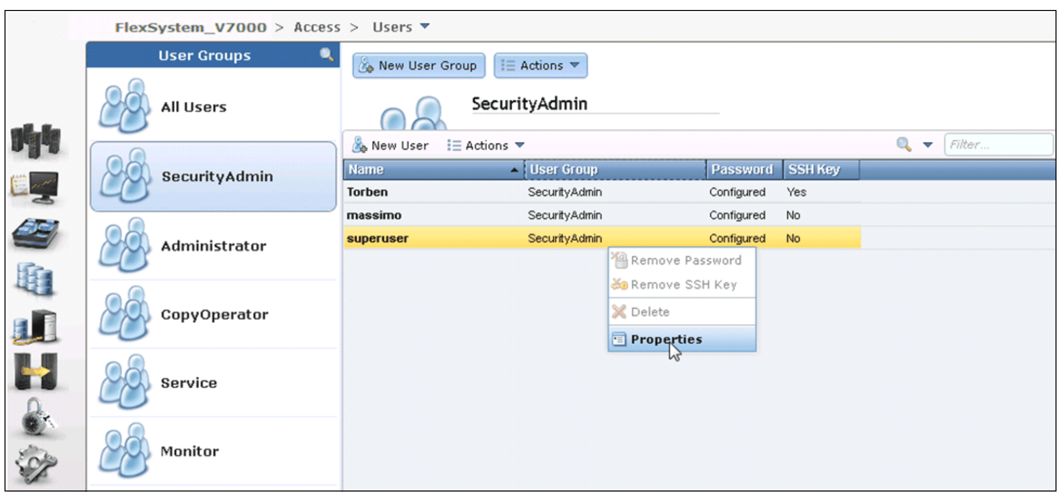


Figure A-6 Superuser properties

3. To upload the public key, click **Browse**, select your public key, and click **OK** (Figure A-7).

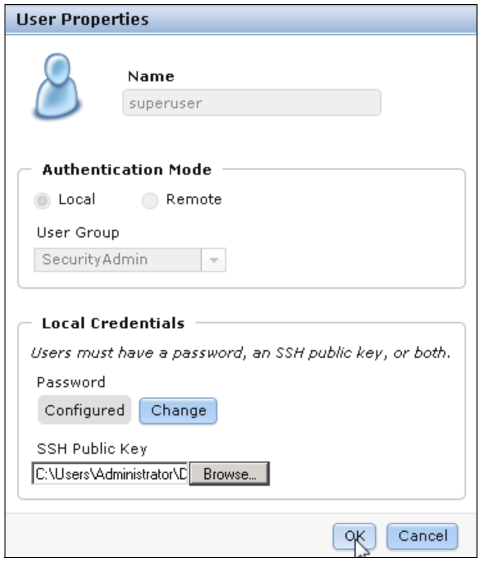


Figure A-7 Select public key

4. Click **OK** and the key is uploaded (Figure A-8).

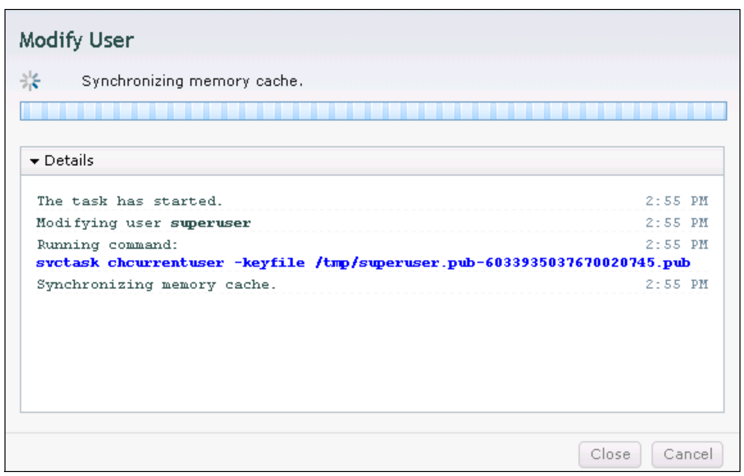


Figure A-8 Public key upload complete

5. Click **Close** to return to the GUI.

Configuring the SSH client

Before the CLI can be used, the SSH client must be configured. To accomplish this task, complete the following steps:

1. Start PuTTY (Figure A-9 on page 580).

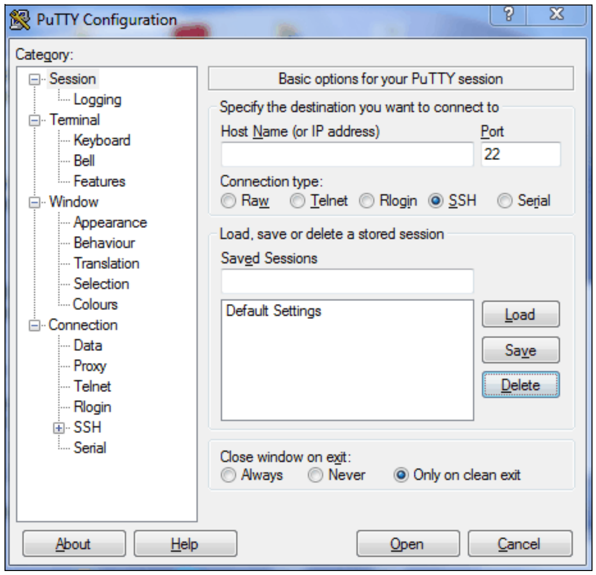


Figure A-9 PuTTY Configuration

In the right pane under the "Specify the destination you want to connect to" section, select **SSH**. Under the "Close window on exit" section, select **Only on clean exit**, which ensures that if there are any connection errors, they are displayed on the user's window.

2. From the Category pane on the left side of the PuTTY Configuration window, click **Connection** → **SSH** to open the PuTTY SSH Configuration window (Figure A-10).

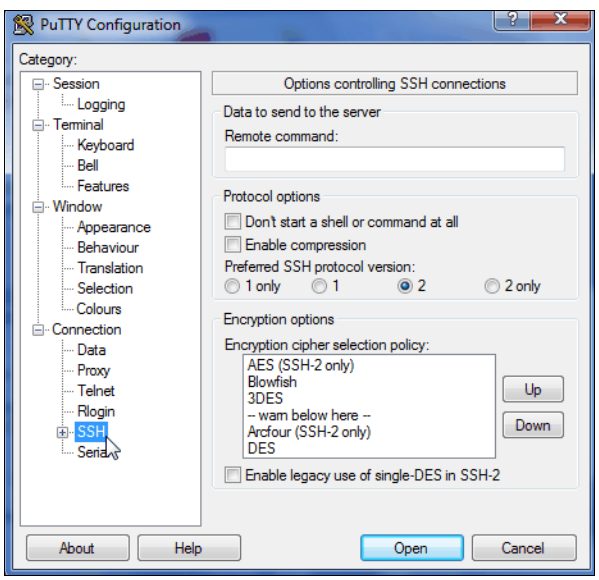


Figure A-10 SSH protocol version 2

3. In the right pane, in the "Preferred SSH protocol version" section, select **2**.

4. From the Category pane on the left side of the PuTTY Configuration window, click **Connection** → **SSH** → **Auth**. As shown in Figure A-11 on page 581, in the right pane, in the "Private key file for authentication:" field under the Authentication Parameters section, either browse to or type the fully qualified directory path and file name of the SSH client private key file created earlier (for example, `C:\Support Utils\PuTTY\icat.PPK`).

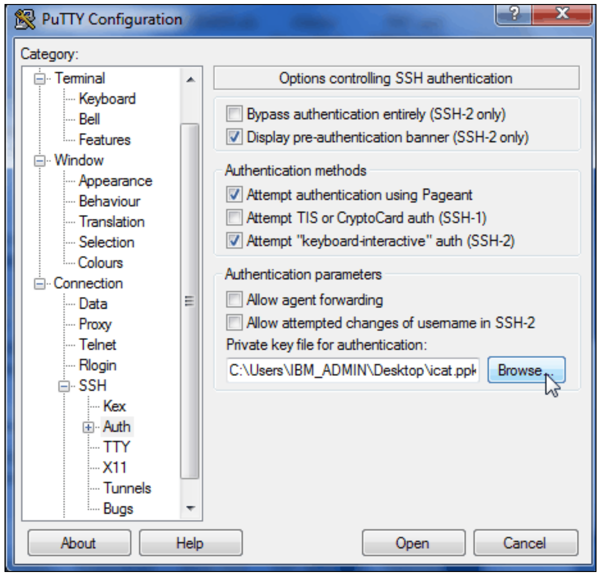


Figure A-11 SSH authentication

5. From the Category pane on the left side of the PuTTY Configuration window, click **Session** to return to the Session view (Figure A-9 on page 580).

6. In the right pane, enter the host name or system IP address of the IBM Flex System V7000 Storage Node cluster in the Host Name field, and enter a session name in the Saved Sessions field (Figure A-12).

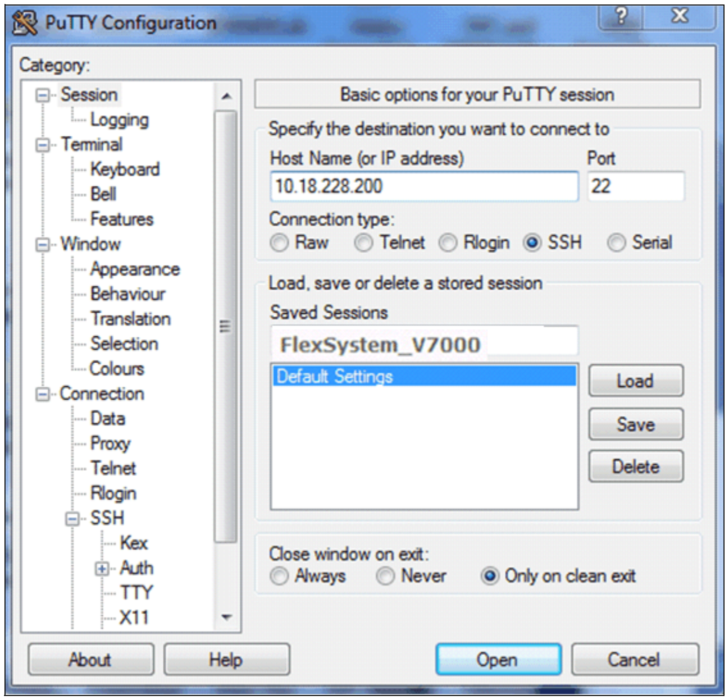


Figure A-12 Enter session information

7. Click **Save** to save the new session (Figure A-13).

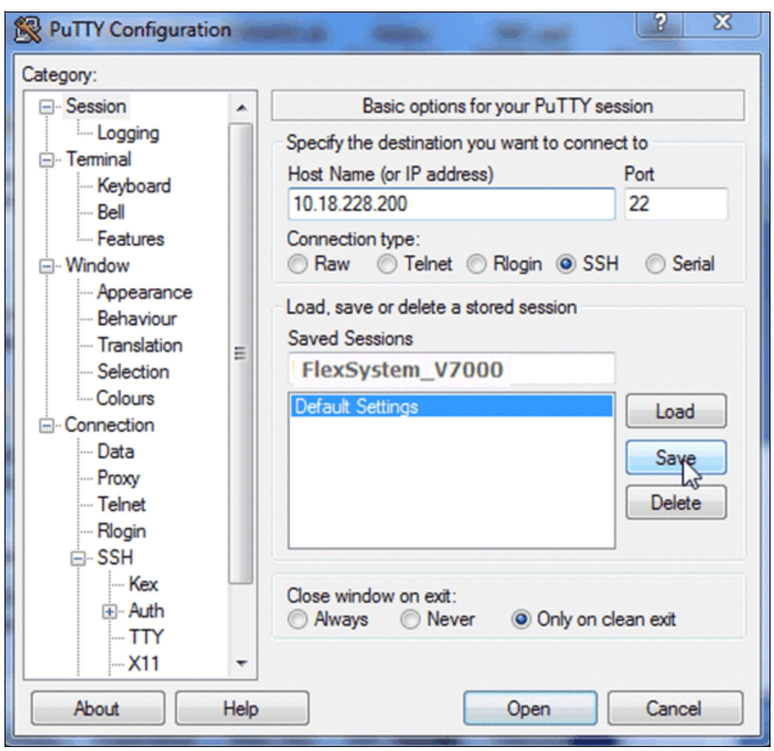


Figure A-13 Save Session

8. Highlight the new session and click **Open** to connect to the IBM Flex System V7000 Storage Node. A PuTTY Security Alert opens; confirm it by clicking **Yes** (Figure A-14).

Figure A-14 Confirm Security Alert

9. PuTTY now connects to the system and prompts you for a user name. Enter admin as the user name and press Enter (Example A-1).

Example: A-1 Enter user name

```
login as: admin
Authenticating with public key "rsa-key-20101006"
IBM_4939:FlexSystem_V7000:admin>
```

You have now completed the tasks that are required to configure the CLI for IBM Flex System V7000 Storage Node administration.

Example commands

A detailed description about all the available commands is beyond the intended scope of this book. In this section are sample commands that we have referenced in this book.

With earlier releases of the IBM SAN Virtual Controller and the IBM Storwize V7000 there were prefixes required to run CLI commands; these were **svcinfo** and **svctask** and are no longer needed in the IBM Flex System V7000 Storage Node. If you have scripts from those platforms that use these prefixes, they run without problems. If you type **svcinfo** or **svctask** and press the Tab key twice, all the available subcommands are listed for those prefixes. Pressing the Tab key twice also auto-completes commands if the input is valid and unique to the system.

Enter **lsvdisk**, as shown in Example A-2, to list all configured volumes on the system. The example shows that three volumes are configured.

Example: A-2 List all volumes

```
IBM_4939:FlexSystem_V7000:admin>lsvdisk
id name             IO_group_id IO_group_name status mdisk_grp_id mdisk_grp_name capacity  type FC_id FC_name RC_id
RC_name vdisk_UID                        fc_map_count copy_count fast_write_state se_copy_count RC_change
0  Hyper-V_PCO_Node1 0          io_grp0       online many         many          100.00GB  many
600507680185853FF00000000000000F 0          2          empty            0             no
1  ESXi-PCO          0          io_grp0       online many         many          100.00GB  many
600507680185853FF000000000000010 0          2          empty            0             no
2  ESXi-Redbooks     0          io_grp0       online many         many          400.00GB  many
600507680185853FF000000000000011 0          2          not_empty        0             no
3  Hyper-V_PCO_Node2 0          io_grp0       online many         many          101.00GB  many
600507680185853FF000000000000012 0          2          empty            0             no
```

```
4  ESXi-Demo        0          io_grp0       online many        many          1000.00GB many
600507680185853FF000000000000013 0             2          not_empty       0             no
```

Enter **lshost** to get a list of all configured hosts on the system (Example A-3).

Example: A-3 List hosts

```
IBM_4939:FlexSystem_V7000:admin>lshost
id name          port_count iogrp_count status
2  HyperV-1_FCoE 2          4           online
3  HyperV-2_FCoE 2          4           online
4  ESXi-1        2          4           online
5  ESXi-2        2          4           online
```

To map the volume to the hosts, enter **mkvdiskhostmap** (Example A-4).

Example: A-4 Map volumes to host

```
IBM_4939:FlexSystem_V7000:admin>mkvdiskhostmap -host ESXi-1 -scsi 0   -force ESXi-Redbooks
Virtual Disk to Host map, id [0], successfully created
```

To verify the host mapping, enter **lsvdiskhostmap** (Example A-5).

Example: A-5 List all hosts mapped to a volume

```
IBM_4939:FlexSystem_V7000:admin>lshostvdiskmap ESXi-1
id name   SCSI_id vdisk_id vdisk_name    vdisk_UID
4  ESXi-1 0       2        ESXi-Redbooks 600507680185853FF000000000000011
```

In the CLI, there are more options available than in the GUI. All advanced settings can be set, for example, I/O throttling. To enable I/O throttling, change the properties of a volume using the **changevdisk** command (Example A-6). To verify the changes, run **lsvdisk**.

> **Command output:** The **lsvdisk** command lists all available properties of a volume and its copies, but to make it easier to read, lines in the example output have been deleted.

Example: A-6 Enable advanced properties: I/O throttling

```
IBM_4939:FlexSystem_V7000:admin>chvdisk -rate 1200 -unit mb ESXi-Redbooks
IBM_4939:FlexSystem_V7000:admin>
IBM_4939:FlexSystem_V7000:admin>lsvdisk ESXi-Redbooks

id 2
name ESXi-Redbooks
.
.
vdisk_UID 600507680185853FF000000000000011
virtual_disk_throttling (MB) 1200
preferred_node_id 2
.
.
IBM_4939:FlexSystem_V7000:admin>
```

If you do not specify the unit parameter, the throttling is based on I/Os instead of throughput (Example A-7 on page 585).

Example: A-7 Throttling based on I/O

```
IBM_4939:FlexSystem_V7000:admin>chvdisk -rate 4000 ESXi-Redbooks
IBM_4939:FlexSystem_V7000:admin>lsvdisk ESXi-Redbooks
id 2
name ESXi-Redbooks
.
.
vdisk_UID 600507680185853FF000000000000011
throttling 4000
preferred_node_id 2
.
.
IBM_4939:FlexSystem_V7000:admin>
```

To disable I/O throttling, set the I/O rate to 0 (Example A-8).

Example: A-8 Disable I/O Throttling

```
IBM_4939:FlexSystem_V7000:admin>chvdisk -rate 0 ESXi-Redbooks
IBM_4939:FlexSystem_V7000:admin>lsvdisk ESXi-Redbooks
id 2
.
.
vdisk_UID 600507680185853FF000000000000011
throttling 0
preferred_node_id 2
.
.
IBM_4939:FlexSystem_V7000:admin>
```

FlashCopy example commands

The IBM Flex System V7000 Storage Node offers powerful FlashCopy functions, as described in Chapter 9, “IBM Flex System V7000 Storage Node Copy Services” on page 341. FlashCopy creation is a process that is often scripted. The commands to create a FlashCopy are easy to discover if you create a FlashCopy in the GUI, and the CLI commands for your system are listed when the creation completes.

Example A-9 shows you the required steps to prepare a reverse FlashCopy, and shows you the FlashCopy command using the “Reverse” option. As you can see at the end of Example A-9, `FCMAP_rev_1` shows a restoring value of `yes` while the FlashCopy mapping is copying. After it has finished copying, the restoring value field will change to `no`.

Example: A-9 Prepare reverse FlashCopy

```
IBM_4939:FlexSystem_V7000:admin>lsvdisk
id name             IO_group_id IO_group_name status mdisk_grp_id mdisk_grp_name capacity  type    FC_id FC_name RC_id
RC_name vdisk_UID                       fc_map_count copy_count fast_write_state se_copy_count RC_change
5  Volume_FC_S       0           io_grp0       online 0            STGPool_0   5.00GB    striped
600507680185853FF000000000000015 0          1          empty            0             no
6  Volume_FC_T       0           io_grp0      online 0            STGPool_0     5.00GB    striped
600507680185853FF000000000000016 0          1          empty            0             no
7  Volume_FC_T1      0           io_grp0      online 0            STGPool_0     5.00GB    striped
600507680185853FF000000000000017 0          1          empty            0             no

IIBM_4939:FlexSystem_V7000:admin>mkfcmap -source Volume_FC_S -target Volume_FC_T_S1 -name FCMAP_1 -copyrate 50
FlashCopy Mapping, id [0], successfully created
```

```
IBM_4939:FlexSystem_V7000:admin>mkfcmap -source Volume_FC_T_S1 -target Volume_FC_S -name FCMAP_rev_1 -copyrate 50
FlashCopy Mapping, id [1], successfully created

IBM_4939:FlexSystem_V7000:admin>mkfcmap -source Volume_FC_T_S1  -target Volume_FC_T1 -name FCMAP_2 -copyrate 50
FlashCopy Mapping, id [2], successfully created

IBM_4939:FlexSystem_V7000:admin>lsfcmap
id name        source_vdisk_id source_vdisk_name target_vdisk_id target_vdisk_name group_id group_name status
progress copy_rate clean_progress incremental partner_FC_id partner_FC_name restoring start_time rc_controlled
0  FCMAP_1     5               Volume_FC_S       8               Volume_FC_T_S1                      idle_or_copied 0
50        100           off         1             FCMAP_rev_1     no                    no
1  FCMAP_rev_1 8               Volume_FC_T_S1    5               Volume_FC_S                         idle_or_copied 0
50        100           off         0             FCMAP_1         no                    no
2  FCMAP_2     8               Volume_FC_T_S1    7               Volume_FC_T1                        idle_or_copied 0
50        100           off                                       no                    no

IBM_4939:FlexSystem_V7000:admin>startfcmap -prep FCMAP_1
IBM_4939:FlexSystem_V7000:admin>lsfcmap
id name        source_vdisk_id source_vdisk_name target_vdisk_id target_vdisk_name group_id group_name status
progress copy_rate clean_progress incremental partner_FC_id partner_FC_name restoring start_time   rc_controlled
0  FCMAP_1     5               Volume_FC_S       8               Volume_FC_T_S1                      copying       0
50        100           off         1             FCMAP_rev_1     no        111006160239 no
1  FCMAP_rev_1 8               Volume_FC_T_S1    5               Volume_FC_S                         idle_or_copied 0
50        100           off         0             FCMAP_1         no                       no
2  FCMAP_2     8               Volume_FC_T_S1    7               Volume_FC_T1                        idle_or_copied 0
50        100           off                                       no                       no

IBM_4939:FlexSystem_V7000:admin>startfcmap -prep FCMAP_2
IBM_4939:FlexSystem_V7000:admin>lsfcmap
id name        source_vdisk_id source_vdisk_name target_vdisk_id target_vdisk_name group_id group_name status
progress copy_rate clean_progress incremental partner_FC_id partner_FC_name restoring start_time   rc_controlled
0  FCMAP_1     5               Volume_FC_S       8               Volume_FC_T_S1                      copying       3
50        96            off         1             FCMAP_rev_1     no        111006160239 no
1  FCMAP_rev_1 8               Volume_FC_T_S1    5               Volume_FC_S                         idle_or_copied 0
50        100           off         0             FCMAP_1         no                       no
2  FCMAP_2     8               Volume_FC_T_S1    7               Volume_FC_T1                        copying       0
50        100           off                                       no        111006160359 no

IBM_4939:FlexSystem_V7000:admin>startfcmap -prep FCMAP_rev_1
CMMVC6298E The command failed because a target VDisk has dependent FlashCopy mappings.
IBM_4939:FlexSystem_V7000:admin>startfcmap -prep -restore FCMAP_rev_1
IBM_4939:FlexSystem_V7000:admin>lsfcmap
id name        source_vdisk_id source_vdisk_name target_vdisk_id target_vdisk_name group_id group_name status  progress
copy_rate clean_progress incremental partner_FC_id partner_FC_name restoring start_time   rc_controlled
0  FCMAP_1     5                 Volume_FC_S       8               Volume_FC_T_S1                          copying 18
50        81            off         1             FCMAP_rev_1     no        111006160239 no
1  FCMAP_rev_1 8                 Volume_FC_T_S1    5               Volume_FC_S                             copying 81
50        18            off         0             FCMAP_1         yes       111006161024 no
2  FCMAP_2     8                 Volume_FC_T_S1    7               Volume_FC_T1                            copying 14
50        100           off                                       no        111006160359 no
```

Modifying Global Mirror Delay for testing

As described in 9.3, “Remote Copy” on page 380, a delay can be simulated on Global Mirror links. This feature is an advanced feature only available in the CLI. Example A-10 shows you how to add a delay to the Global Mirror connection.

Example: A-10 Add Global Mirror Delay

```
IBM_4939:FlexSystem_V7000:admin>lssystem
id 0000020061614FFC
name FlexSystem_V7000
.
console_IP 10.18.228.200:443
```

```
id_alias 0000020061614FFC
gm_link_tolerance 300
gm_inter_cluster_delay_simulation 0
gm_intra_cluster_delay_simulation 0
gm_max_host_delay 5
.
tier generic_hdd
tier_capacity 5.45TB
tier_free_capacity 2.10TB
has_nas_key no
layer storage
rc_buffer_size 48

IBM_4939:FlexSystem_V7000:admin>lssystem
id 0000020061614FFC
name FlexSystem_V7000
gm_inter_cluster_delay_simulation 0
gm_intra_cluster_delay_simulation 20
gm_max_host_delay 5
.
IBM_4939:FlexSystem_V7000:admin>chsystem -gminterdelaysimulation 20
IBM_4939:FlexSystem_V7000:admin>lssystem
id 0000020061614FFC
name FlexSystem_V7000
.
gm_inter_cluster_delay_simulation 20
gm_intra_cluster_delay_simulation 20
gm_max_host_delay 5
.
```

For more details on the various commands and their usage, as well as the use of the command line interface itself; refer to *Command-Line Interface User's Guide GC27-2287-03* located in the Info Center of the IBM Flex System V7000 Storage Node at:

http://publib.boulder.ibm.com/infocenter/flexsys/information/index.jsp

Author comment - This is where the FlexV700 GUI sends us; need to watch for an update to be added here or the link to change. Also need to watch for any possible document # changes as this is the V7000 version for V6.4.0 release...

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *Implementing the IBM System Storage SAN Volume Controller V6.3*, SG24-7933
- *Implementing the IBM Storwize V7000 V6.3*, SG24-7938
- *IBM PureFlex System and IBM Flex System Products and Technology*, SG24-7984
- *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859
- *Implementing Systems Management of IBM PureFlex System*, SG24-8060

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

These publications are also relevant as further information sources:

- IBM PureSystems Redbooks pages:

  http://www.redbooks.ibm.com/portals/puresystems
- IBM Storwize V7000 and IBM Flex System Redbooks pages:

  http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=v7000
- IBM Flex System Information Center:

  http://publib.boulder.ibm.com/infocenter/flexsys/information/index.jsp
- IBM Flex System V7000 Storage Node Information Center:

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_product_page.html
- IBM Flex System V7000 Storage Node Installation Guide (GC27-4203)

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_bkmap_quickinstbk.pdf
- IBM Storwize V7000 Expansion Enclosure Installation Guide, Machine type 2076 (GC27-4234)

  http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/tbrd_bkmap_install_2076.pdf
- IBM Flex System V7000 Storage Node Troubleshooting, Recovery, and Maintenance Guide (GC27-4205)

http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.doc/site_bkmap_trblshoot.pdf

Online resources

These websites are also relevant as further information sources:

- The IBM Storwize V7000 Support page is at the following website:

  http://www.ibm.com/storage/support/storwize/v7000
- The IBM Storwize V7000 home page is at the following website:

  http://www-03.ibm.com/systems/storage/news/center/storwize_v7000/index.html
- The IBM Storwize V7000 Supported hardware list:

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003908
- The IBM Storwize V7000 Configuration Limit and Restrictions are at the following website:

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003902
- IBM Flex System V7000 Storage Node - Learning Modules
- http://publib.boulder.ibm.com/infocenter/flexsys/information/topic/com.ibm.acc.4939.learning/site_learning_tutorials.html
- You can see IBM Redbooks publications about both the IBM Flex System and IBM Storwize V7000 Storage at the following website:

  http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=v7000
- The IBM Storwize V7000 Storage Support page is at the following website

  http://www.ibm.com/storage/support/storwize/v7000
- The IBM Storwize V7000 Storage home page is at the following website
  http://www-03.ibm.com/systems/storage/news/center/storwize_v7000/index.html
- The Storwize V7000 Storage Supported hardware list is at the following website

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003908
- The Storwize V7000 Storage Configuration Limit and Restrictions are at the following website

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003902
- The Storwize V7000 Storage Online documentation is at the following website

  http://www-947.ibm.com/support/entry/portal/Documentation/Hardware/System_Storage/Disk_systems/Mid-range_disk_systems/IBM_Storwize_V7000_%282076%29
- The Storwize V7000 Storage Online Information Center is at the following website:

  http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(-->Hide:)>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine.fm still open and **File>Import>Formats** the Conditional Text Settings (ONLY!) to the book files.

IBM Flex System V7000 Storage Node Introduction and Implementation Guide

(1.5” spine)
1.5”<-> 1.998”
789 <->1051 pages

IBM Flex System V7000 Storage Node Introduction and Implementation

(1.0” spine)
0.875”<->1.498”
460 <-> 788 pages

IBM Flex System V7000 Storage Node Introduction and Implementation

(0.5” spine)
0.475”<->0.873”
250 <-> 459 pages

IBM Flex System V7000 Storage Node Introduction and Implementation Guide

(0.2”spine)
0.17”<->0.473”
90<->249 pages

(0.1”spine)
0.1”<->0.169”
53<->89 pages

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(-->Hide:)>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine.fm still open and **File>Import>Formats** the Conditional Text Settings (ONLY!) to the book files.

IBM Redbooks

IBM Flex System V7000 Storage Node Introduction and Implementation Guide

(2.5" spine)
2.5"<->nnn.n"
1315<-> nnnn pages

IBM Redbooks

IBM Flex System V7000 Storage Node Introduction and Implementation Guide

(2.0" spine)
2.0" <-> 2.498"
1052 <-> 1314 pages

Draft Document for Review January 29, 2013 12:52 pm

IBM Flex System V7000 Storage Node Introduction and Implementation Guide

Introduction to IBM Flex System Family

IBM Flex System V7000 Storage Node Hardware Introduction

Host Configuration Guide

IBM® Flex System™ products are ideally suited for data center environments that require flexible, cost-effective, secure, and energy-efficient hardware.The IBM Flex System V7000 Storage Node is the latest addition to the IBM Flex Systems product family and is a modular storage system designed to fit into the IBM Flex System Enterprise chassis.

When purchased in the IBM PureFlex™ configurations, the IBM Flex System V7000 Storage Node is configured from the factory into the hardware solution purchased. If however, the desired configuration is not offered in the predefined offerings then a "Build to Order" configuration is designed to meet your needs.

The IBM Flex System V7000 Storage Node includes the capability to virtualize its own internal storage in the same manner as the IBM Storwize® V7000 does, and is designed to be a scalable internal storage system to support the compute nodes of the IBM Flex System environment.

This IBM Redbooks® publication introduces the features and functions of the IBM Flex System V7000 Storage Node through several examples. This book is aimed at pre- and post-sales technical support and marketing, storage administrators, and will help you understand the architecture of the IBM Flex System V7000 Storage Node, how to implement it, and take advantage of the industry leading functions and features

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-8068-00 ISBN